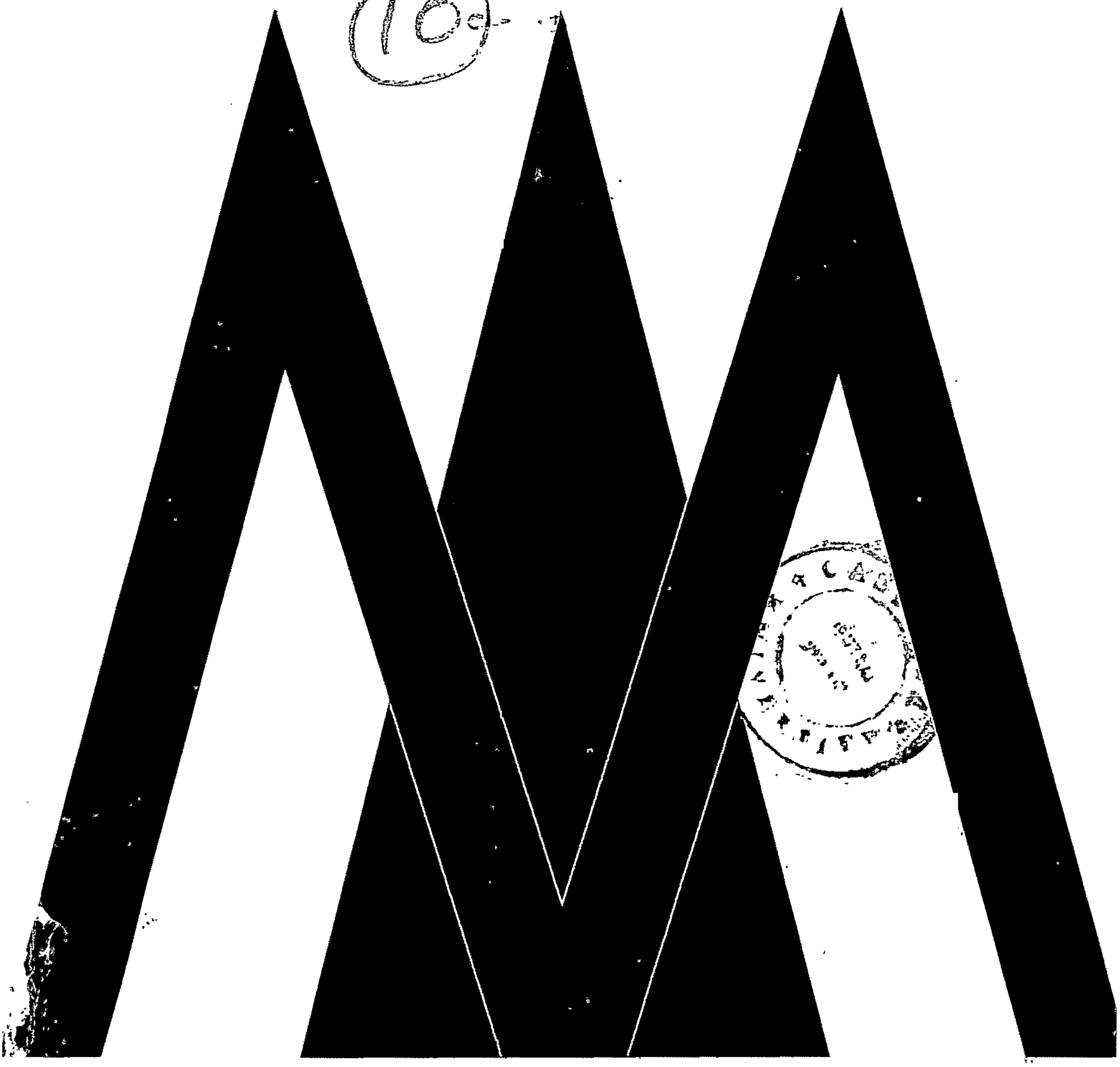
Cub. 11. 2060-10. 1133735

Academy of Management
JOURNAL

658.005
Ae 21

100-



VOLUME 29 NUMBER 1 MARCH 198

ACADEMY OF MANAGEMENT

BOARD OF GOVERNORS

PRESIDENT	Fred Luthans, *University of Nebraska*
PRESIDENT-ELECT & PROFESSIONAL DIVISIONS COORDINATOR	Richard M. Steers, *University of Oregon*
VICE PRESIDENT & PROGRAM CHAIRPERSON	Don Hellriegel, *Texas A&M University*
SECRETARY-TREASURER & BUSINESS MANAGER, ACADEMY OF MANAGEMENT PUBLICATIONS	Walter B. Newsom, *Mississippi State University*
PAST PRESIDENT	Kathryn M. Bartol, *University of Maryland*
EDITOR, ACADEMY OF MANAGEMENT EXECUTIVE	W. Warner Burke, *Columbia University*
EDITOR, ACADEMY OF MANAGEMENT JOURNAL	Janice M. Beyer, *State University of New York-Buffalo*
EDITOR, ACADEMY OF MANAGEMENT REVIEW	Orlando Behling, *Bowling Green State University*
DIRECTOR OF MEMBERSHIP	Joseph P. Yaney, *Texas Tech University*
REPRESENTATIVES-AT-LARGE	Joseph A. Alutto, *State University of New York-Buffalo* Michael Beer, *Harvard Business School* Richard B. Chase, *University of Southern California* John M. Ivancevich, *University of Houston* Edwin L. Miller, *University of Michigan* Jeffrey Pfeffer, *Stanford University* Karlene H. Roberts, *University of California-Berkeley* Mary Ann Von Glinow, *University of Southern California*

demy of Manage

CHAIRPERSONS OF PROFESSIONAL DIVISIONS

BUSINESS POLICY & PLANNING	David B. Jemison, *Stanford University*
CAREERS	Donald D. Bowen, *University of Tulsa*
HEALTH CARE ADMINISTRATION	Michael Abelson, *Texas A&M University*
INTERNATIONAL MANAGEMENT	Bhal Bhatt, *University of Toledo*
MANAGEMENT EDUCATION & DEVELOPMENT	William J. Heisler, *Newport News Shipbuilding*
MANAGEMENT HISTORY	Mariann Jelinek, *Case Western Reserve University*
MANAGERIAL CONSULTATION	Linda M. Calvert, *University of Houston-Clear Lake*
ORGANIZATION & MANAGEMENT THEORY	Gerald R. Salancik, *University of Illinois*
ORGANIZATIONAL BEHAVIOR	Randall B. Dunham, *University of Wisconsin*
ORGANIZATIONAL COMMUNICATION	Susan A. Hellweg, *San Diego State University*
ORGANIZATIONAL DEVELOPMENT	Marshall Sashkin, *National Institute of Education*
PERSONNEL/HUMAN RESOURCES	Marc Wallace, *University of Kentucky*
PRODUCTION/OPERATIONS MANAGEMENT	Kathleen A. Johnson, *University of Louisville*
PUBLIC SECTOR	Barry Bozeman, *Syracuse University*
SOCIAL ISSUES IN MANAGEMENT	Frederick D. Sturdivant, *Ohio State University*
WOMEN IN MANAGEMENT	V. Jean Ramsey, *Western Michigan University*

OTHER OPERATIONS

ACADEMY HISTORIAN & ARCHIVIST	Charles D. Wrege, *Rutgers University*
NEWSLETTER EDITOR	Rose Knotts, *North Texas State University*
PLACEMENT CHAIRPERSON	James B. Townsend, *Kansas State University*
SIGMA IOTA EPSILON	B.J. Hodge, *Florida State University*
PROCEEDINGS EDITORS	John A. Pearce II and Richard B. Robinson, Jr., *University of South Carolina*

P 3955

Academy of Management JOURNAL

10

EDITOR Janice M. Beyer, *State University of New York at Buffalo*
CONSULTING EDITORS Richard Mowday, *University of Oregon*
Manuel London, *AT&T Communications*
Carl Zeithaml, *Texas A&M University*
EDITORIAL BOARD Harold L. Angle, *University of Minnesota*
Achilles A. Armenakis, *Auburn University*
Richard Arvey, *University of Minnesota*
Susan Ashford, *Dartmouth College*
Jean M. Bartunek, *Boston College*
Barry Baysinger, *Texas A&M University*
Max H. Bazerman, *Massachusetts Institute of Technology*
Richard S. Blackburn, *University of North Carolina*
Joel Brockner, *Columbia University*
Terry Connolly, *University of Arizona*
Robert D. Dewar, *Northwestern University*
Jane E. Dutton, *New York University*
Cecilia M. Falbe, *Baruch College, CUNY*
Daniel C. Feldman, *University of Florida*
Peter Feuille, *University of Illinois*
Charles Fombrun, *University of Pennsylvania*
Jeffrey D. Ford, *Ohio State University*
Cynthia V. Fukami, *University of Denver*
William Glick, *University of Texas*
Robert Golembiewski, *University of Georgia*
Donald C. Hambrick, *Columbia University*
Kathryn R. Harrigan, *Columbia University*
Frederic M. Jablin, *University of Texas*
William Joyce, *Dartmouth College*
Richard J. Klimoski, *Ohio State University*
Thomas Kochan, *Massachusetts Institute of Technology*
Linda A. Krefting, *Texas Tech University*
Frank Krzystofiak, *State University of New York at Buffalo*
Gary P. Latham, *University of Washington*
Alan D. Meyer, *University of Oregon*
Peter Monge, *University of Southern California*
Cynthia A. Montgomery, *Northwestern University*
Jerry M. Newman, *State University of New York at Buffalo*
John Newstrom, *University of Minnesota-Duluth*
Jon L. Pierce, *University of Minnesota-Duluth*
Jerry Porras, *Stanford University*
Claudia B. Schoonhoven, *San Jose State University*
John E. Sheridan, *Texas Christian University*
Harbir Singh, *University of Pennsylvania*
James R. Terborg, *University of Oregon*
Dave O. Ulrich, *University of Michigan*
Alan R. Wilkins, *Brigham Young University*
Gary A. Yukl, *State University of New York at Albany*
Ray Zammuto, NCHEMS; *University of Colorado-Denver*
Dale Zand, *New York University*
INDEX EDITOR Nancy G. Dodd, *University of Hawaii-Hilo*
BUSINESS MANAGER Walter B. Newsom, *Mississippi State University*
ADVERTISING DIRECTOR Donna Ledgerwood, *North Texas State University*
PRODUCTION EDITOR Steven Gomes, *Buffalo, New York*
COPY EDITOR Seph Weene, *Ithaca, New York*
EDITORIAL ASSISTANT Denise Fredericks

P3955

INFORMATION FOR CONTRIBUTORS

The *Academy of Management Journal* publishes articles in fields of interest to members of the Academy of Management. These fields of interest are reflected in the divisions and interest groups of the Academy listed on the inside front cover.

The *Journal* publishes only original, empirical research as articles or research notes. Manuscripts are considered with the understanding that their contents have not been published and are not under consideration elsewhere. In its articles, the *Journal* seeks to publish reports of research that develops, tests, or advances management theory and practice. All types of empirical methods—quantitative, qualitative, or combinations—are acceptable. The *Journal* does not publish purely conceptual or review articles; these are published in the *Academy of Management Review*. Atheoretical exploratory or survey research, methodological studies, replications or extensions of past research, and commentaries with new empirical content are also of interest for publication as research notes if they make an important contribution to knowledge relevant to management.

Articles and research notes should be written so they are understandable and interesting to all members of the Academy. The contributions of research in all specialized areas to *general* management theory and practice should be made evident. Specialized argot and jargon should be translated into terminology in general use within the fields of management.

Articles should also be written as concisely as possible without sacrificing meaningfulness or clarity of presentation. To save space, tables should be combined and data should be presented in the text wherever possible. Manuscripts submitted for publication as articles should not ordinarily exceed 30 double-spaced typewritten pages, including tables. Manuscripts submitted for publication as research notes should not exceed ten double-spaced typewritten pages, including tables. Everything, including tables, in submitted manuscripts should be typed in double-spaced format on one side of the page. Manuscripts prepared on computers should be printed on letter-quality printers or, if other printers are used, in double-strike or enhanced print.

Decisions regarding publication of submitted manuscripts are based upon the recommendations of members of the Editorial Review Board and other qualified reviewers. All articles and research notes published in the *Journal* are subjected to the same anonymous review process. Reviewers evaluate manuscripts on their significance to the field, their conceptual adequacy, their technical adequacy, their appropriateness, and their clarity of presentation. Reviewers' comments are made available to authors.

Submissions should be sent to Dr. Janice M. Beyer, Editor, *Academy of Management Journal*, School of Management, Jacobs Management Center, State University of New York, Buffalo, New York 14260. Manuscripts should be prepared in accordance with the *AMJ Style Guide for Authors*, which is published periodically and is also available from the editor. Contributors should submit four copies of papers, retaining the original for their files. The *Journal* does not return manuscripts unless requested to do so.

Academy of Management JOURNAL

CONTENTS

Volume 29, Number 1, March 1986

Functions of Parting Ceremonies in Dying Organizations
Stanley G. Harris and Robert I. Sutton 5

Dual Commitment and Labor-Management Relationship Climates
Harold L. Angle and James L. Perry 31

Group Approaches for Improving Strategic Decision Making: A Comparative Analysis of Dialetical Inquiry, Devil's Advocacy, and Consensus
David M. Schweiger, William R. Sandberg, and James W. Ragan 51

Administrative Succession and Organizational Performance: How Administrator Experience Mediates the Succession Effect
Jeffrey Pfeffer and Alison Davis-Blake 72

Effects of Alternative Job Evaluation Methods on Decisions Involving Pay Equity
Robert M. Madigan and David J. Hoover 84

The Relationship of Seniority to Job Performance Following Reinstatement
Robert C. Rodgers, I. B. Helburn, and John E. Hunter 101

Effects of Gender on Self- and Supervisory Ratings
Lynn McFarlane Shore and George C. Thornton III 115

RESEARCH NOTES

Using the Lens Model to Magnify Raters' Consistency, Matching, and Shared Bias
Neal Schmitt, Raymond A. Noe, and Rand Gottschalk 130

Effects of Task Difficulty and Interunit Interdependence on Information Processing Systems
Jack K. Ito and Richard B. Peterson 139

Solution Diversity and the Quality of Group Decisions
John P. Wanous and Margaret A. Youtz 149

Using Electronic Mail for Data Collection in Organizational Research
Lee S. Sproull 159

Attitudes Toward Work Schedules: Construct Definition, Instrument Development, and Validation
Randall B. Dunham and Jon L. Pierce 170

Style Guide For Authors 183

Published by the Academy of Management four times a year in March, June, September, and December. Subscription rates for U.S., Mexico, and Canada: $38 for one year, $70 for two years, $96 for three years. Rates for overseas: $42 for one year, $78 for two years, $108 for three years. All payable in U.S. dollars.

Subscriptions and orders for back issues should be sent to the Business Manager, Dr. Walter B. Newsom, P.O. Drawer KZ, Mississippi State University, Mississippi State, Mississippi 39762.

Manuscripts and other material of an editorial nature should be sent to the Editor, Dr. Janice M. Beyer, School of Management, State University of New York—Buffalo, Buffalo, New York 14260.

Copies of articles published in the *Journal* may be made for instructional, non-commercial use. The Academy of Management does not require prior clearance and makes no charge for this permission.

Inquiries related to advertising should be addressed to Donna Ledgerwood, College of Business Administration, North Texas State University, Denton, Texas, 76203.

Typesetting by Repro Services, Inc., Decatur, Georgia; presswork, binding, and mailing by Darby Printing Company, Atlanta, Georgia.

Academy of Management Journal, Publication No. 900920.

Second class postage paid at Mississippi State, MS, and additional mailing offices.

POSTMASTER—Send Form 3579 to Academy of Management *Journal*, P.O. Drawer KZ, Mississippi State University, Mississippi State, MS 39762.

ISSN, 0001-4273

© *Academy of Management Journal*
1986, Vol. 29, No. 1, 5–30.

FUNCTIONS OF PARTING CEREMONIES IN DYING ORGANIZATIONS

STANLEY G. HARRIS
University of Michigan
ROBERT I. SUTTON
Stanford University

This paper advances a theory about the functions that parting ceremonies serve for members displaced by organizational deaths. The theory is grounded in data from 11 parties, picnics, and dinners that occurred in six dying organizations. We propose that parting ceremonies create settings in which people can exchange emotional support and can edit self-, social system, and event schemata. We also assert that such gatherings provide opportunities for managers to influence the course of organizational demise.

"The wake is to be an occasion to remember the vigor and charm of the departing spirit. We are interested in recalling and sharing memories of the place with those who were its friends and who may have benefited from their association with it over the years. If you cannot attend, please consider sending a message perhaps containing an anecdote you remember with pleasure."
—An invitation to a "wake" for a dying organization.

This article describes the attributes and functions of a set of ceremonies that we observed in an investigation of dying organizations.[1] We were struck by the prevalence of parties, picnics, and other social occasions during the final phases of organizational death. We labeled these events *parting ceremonies*.

Ceremonies are "relatively elaborate, dramatic, planned sets of activities that consolidate various forms of cultural expressions into one event, which is carried out through social interactions" (Trice & Beyer, 1984: 655). Turner (1977) asserted that such ritualistic behaviors arise when the structure of

---

An earlier version of this manuscript was presented at the 1984 meeting of the Academy of Management as part of the "Sense-making in Organizations" symposium co-sponsored by the Organization and Management Theory and the Organizational Behavior divisions. We wish to thank Daniel Denison, Nancy Harris, Robert Kahn, and Karl Weick for their helpful comments on this manuscript and Reginald Bruce and Susan Schurman for their help in collecting these data. We would also like to thank the people who took the time to talk with us about their experiences with organizational death.

[1] The use of the death metaphor is intentional and serves to remind us that organizations are living and dying systems (Miller, 1978) and that there is a strong similarity between the losses people experience at the death of an organization and of a friend. The analysis we present in the following pages has benefited from the mutation of the death metaphor (Weick, 1979). See Sutton (1984) for further discussion of this metaphor.

social relations is threatened. Organizational death causes members and other participants to lose a shared social arena and thus poses a profound threat to the structure of social relations. As a result, an array of rituals and ceremonies characterize dying organizations, including collective bargaining, disposing of property, making public announcements, and severing ties with clients and suppliers. We seek here to propose a theory about one of these activities—the parting ceremony.

Parting ceremonies are formal social events that arise in response to the imminent breaking of social bonds. Like other rites of passage (Van Gennep, 1960), parting ceremonies help separate members from dying organizations and to a lesser extent facilitate their transition and integration into new roles. The stated purpose of such gatherings is usually to offer a final opportunity for everyone affected to meet as a group. Examples of parting ceremonies outside the domain of organizational death include high school graduation exercises, retirement parties, funerals, wakes, and going-away parties.

Parting during organizational death can be accomplished through a variety of socially approved activities. Such parting activities vary on at least two dimensions—formality and number of participants. For example, partings that occur between two friends are usually informal. In contrast, the parting ceremonies described here are relatively formal because someone plans and sponsors them, and they include more people.

Parting ceremonies appear to be prevalent in organizational death; they were a dominant theme in six of the eight closings we studied. Nonetheless, the modest literature on ceremonies in organizations (e.g., Deal & Kennedy, 1982; Trice, 1985; Trice, Belasco, & Alutto, 1969; Trice & Beyer, 1984) does not address parting ceremonies. Trice and Beyer observed that no rites or ceremonials were "found at the organizational level that have the manifest intent of dissolving existing social arrangements without degrading their occupants" (1984: 667). This study offers evidence of the existence of such ceremonies.

We sought to develop a theory about parting ceremonies because they appear to serve important functions during organizational death. Moreover, we found parting ceremonies to be among the most interesting aspects of organizational death. As Davis (1971) made clear, interest may be the most important quality that a theory can possess.

Studying such parting ceremonies also enhances knowledge about the process of organizational death. Death appears to be a common occurrence in all populations of organizations. Starbuck and Nystrom (1981), for example, presented evidence that 45 percent of American corporations die within 5 years of birth, and 90 percent die within 20 years. They also reported that federal agencies die at a similar rate.

Organizational theorists offer an array of causal explanations for this common transformation. The literatures on organizational effectiveness (Cameron & Whetten, 1982; Goodman, Pennings, & Associates, 1977), open systems (Katz & Kahn, 1978), and population ecology (Hannan & Freeman, 1984; McKelvey, 1982), for example, predict when and explain why the event of organizational death occurs. In contrast, little has been written about the process of organizational death. Although some literature has explored conceptual issues that arise when industrial plants close (Hardy, 1982; Taber, Walsh, & Cooke, 1979), most existing descriptions of organizational death are journalistic rather than theoretical (e.g., Nance, 1984; Slote, 1969). Parting ceremonies provide a useful vantage point from which to study this ubiquitous social process.

The theory proposed here emphasizes the positive consequences of parting ceremonies; thus it is in the functional tradition (Trice, 1985). We view organizations as living and dying systems, which is also consistent with the organismic metaphor characteristic of the functional approach (Miller, 1985; Morgan, Frost, & Pondy, 1983). The functional perspective suggests that patterned behavior contributes to the maintenance and the survival of social systems (Trice, 1985). Specifically, the present theory-building effort considers several functions that parting ceremonies serve. These include maintaining and promoting (1) individual well-being, (2) group cohesion, (3) task performance, and (4) positive relations with external stakeholders.

We sought data about both the functions and dysfunctions of parting ceremonies, but the theory proposed here emphasizes the benefits rather than the hazards of such ceremonies. We did not incorporate the dysfunctions of parting ceremonies into the theory because our data and other descriptions of parting ceremonies (Slote, 1969) suggest that such social gatherings have primarily functional consequences.

We acknowledge that we encountered some evidence of dysfunctions. Moreover, other behaviors displayed during parting ceremonies appear to serve no function at all. The neglect of dysfunctions does not invalidate the theory offered here, because the presence of functional consequences is not logically incompatible with the presence of dysfunctional consequences. Yet such neglect does render the theory less than complete. As Weick (1979) pointed out, all organizational theories suffer from this malady because of the complexity of social phenomena.

In their codification of rites, Trice and Beyer (1984) distinguished between those with manifest consequences for individuals and those with manifest consequences for groups or organizations. Similarly, we separately discuss functions that parting ceremonies serve for (1) individuals, (2) organizations, and (3) external stakeholders. However, the theory offered here emphasizes the functions that parting ceremonies serve for displaced organizational members.

Our research was inductive. Following the logic of induction, we describe the research first and propose our theory second.

METHODS

Organizations

This research was conducted in southeastern Michigan from the summer of 1982 through late spring of 1983. Michigan was experiencing a deep economic recession during this period and thus served as a natural laboratory for studying organizational death. The investigation of parting ceremonies described here was part of a larger study about the process of organizational death (Sutton, 1984).

The sampling strategy for this study followed the example of Lawrence and Lorsch's (1967) classic research on environment and organization (Lawrence, 1981). We deliberately, but not randomly, selected a diverse sample of organizational deaths. Such variation was needed to help generate theory about both constant elements that arise in all or almost all organizational deaths and variable elements that explain differences among organizational death processes (Sutton, 1984).

Our efforts to develop theory about parting ceremonies focused on only constant elements. Nonetheless, the varied sample served to strengthen theory-building about constant elements that we proposed are present in all, or nearly all, parting ceremonies. Similarities observed across a diverse sample offer firmer grounding for such propositions than constant elements observed in a homogeneous sample of organizations. We sought a sample of organizational deaths that would fill four theoretically derived categories (see Table 1).

The final sample included the following: in category 1, private sector organizations with no parent organizations, a small automobile company

TABLE 1
Core Cases of Organizational Death

	Organization	Size[a]	Age (Years)	Informants[c]
Private sector				
Independent	Small automobile company	30 1	5½	2
	Sporting goods store	4 3	1	1
Dependent on a parent organization	Hospital	180 135	46	8
	Retail store	? ?	9	4
	Large auto plant	550 10	25	8
Public sector				
Freestanding facility	Center for the disabled	374 13	20	8
Part of a larger facility	Academic unit	14[b] 13	58	8
	Research organization	30 18	16	7

[a] The top number indicates how many employees worked for the defunct organization one year prior to its closing; the bottom figure indicates the number one month before the closing. These data are missing for the large department store.

[b] These numbers refer to faculty members only; support staff and graduate students are excluded.

[c] Note that while the total here is 46 informants, this actually represents 44 individuals. Two individuals, members of a shared parent organization, acted as decision makers for two distinct closings.

and a small sporting goods store; in category 2, private sector organizations dependent on a parent organization, a hospital, a large retail store, and a large auto plant; in category 3, public sector, freestanding facilities, a center for the disabled; and in category 4, public sector organizations embedded in larger facilities, an academic unit and a research organization.

In addition to variation in technology, sector, and dependence, Table 1 reveals much variation in organizational age (1 to 58 years) and number of members (4 to 550) within this small sample.

Three additional criteria guided the selection of cases. First, organizations had to be in the process of closing or have been closed no more than one year before data gathering began. Second, we had to be allowed to study the death. Management had to grant permission if an organization were still alive. For organizations that were defunct, we contacted displaced members. Cases were included if enough key individuals were willing to discuss a closing. We contacted members of 20 organizations before reaching our target, a sample of eight organizational deaths.

Finally, informants had to agree that there was no chance that their organizations would start activities again in the future. Following living systems theory (Katz & Kahn, 1978; Miller, 1978), we operationally defined organizations as dead when their input-throughput-output cycles had ceased permanently.

Informants

People who made and implemented decisions about an organization's closing were our primary informants. We selected and tried to interview the head of each organization and other key decision makers including human resource managers, marketing managers, and operations officers. We also sought to interview decision makers from parent organizations and union leaders when such individuals existed. Moreover, we interviewed informants who were not decision makers when they had essential information.

The perspectives of decision makers were emphasized because they usually have high network centrality (Tichy, 1981: 229). Compared to other members, decision makers typically have structural connections to the greatest numbers of people in organizations, receive the most information, and send the most information. They are also most likely to have access to organizational records. Moreover, they are usually the last people to be displaced. Thus, decision makers are most likely to have observed all phases of organizational death, including parting ceremonies.

The number of informants interviewed for each closing (see Table 1) followed the notion of "theoretical saturation" suggested by Glaser and Strauss (1967: 61–62); we stopped interviewing additional people when we began hearing the same stories repeated over and over. We invited 55 people to participate in the study. A final sample of 44 agreed to participate. Of these 44 informants, 26 were members of the dying organizations studied and 18 were nonmembers. Nonmembers included decision makers from parent organizations and union leaders.

Data Sources

The case descriptions of the eight organizational deaths are based primarily on responses to semistructured interviews and on archival data. Each interview began with questions about the norms and history of an informant's organization. The next questions concerned the early stages of the organization's death including rumors, the struggle to save the organization, the decision to close it, and the announcement of the closing. Next were questions about the events during the process of disbanding the organization.

The remainder of the interview concerned the final days and hours of the organizational closing, and the psychological and economic effects of the closing on various stakeholders. This section included questions about the existence and occurrence of parting ceremonies. Moreover, since this research was inductive, we encouraged interviewers to ask questions about lines of discussion that deviated from the prepared questions.

Most interviews (40 of 44) were tape-recorded and transcribed. The length of these contrived conversations varied from 45 minutes to over four hours; a typical interview was about 90 minutes. Interviews usually took only one session, but occasionally researchers or informants requested more sessions. For example, we conducted an additional interview with the head of the center for the disabled to learn more about a trio of parting ceremonies.

Access to records produced during the organizational deaths was considered essential. We asked each informant if any records existed about a closing, where they were kept, and if we could have permission to copy the documents. We gathered useful data from such records as letters, lay-off lists, budgets, and newspaper stories for all the organizations except the small sporting goods store. For instance, the quote that introduces this paper came from an invitation to a "wake" for the academic department.

Analysis Strategy

The method used to develop theory about parting ceremonies drew on existing descriptions of how to generate theory from qualitative data, particularly writings by Glaser and Strauss (1967), Mintzberg (1979), and Miles (1979). In general, our method entailed continually comparing theory and data until we had developed adequate conceptual categories.

Specifically, efforts to develop a theory about parting ceremonies went on within the larger study of which this research was a part. Before and during data collection, we met every week to discuss the emerging theory. Our earliest meetings focused on constructing a rough working framework. We heeded Miles's (1979) warning that "the need to develop grounded theory exists in tension with the need for clarity and focus; research projects that pretend to come to the study with no assumptions usually encounter much difficulty" (1979: 119).

Subsequent meetings entailed comparing the emerging framework to new evidence from the field study and the organizational studies literature; adjusting the framework to allow for new facts and new ideas; and planning to collect additional evidence. One week, for example, we noted that former members of dying organizations usually receive invitations to parting ceremonies. We developed hypotheses about the functions such invitations serve. Plans were made to call back an informant and ask him if former members of his defunct organization had been invited to parting ceremonies. Finally, we decided to urge the other two members of the research team to ask informants about such invitations.

This iterative approach is similar to the constant comparison method suggested by Glaser and Strauss (1967). We used this method to generate a theory about the functions that parting ceremonies serve for displaced members and to identify functions that such ceremonies serve for the management of dying organizations and their parent organizations.

THE FUNCTIONS PARTING CEREMONIES SERVE FOR DISPLACED MEMBERS

The members of dying organizations share a variety of fates. Organizational demise—in contrast to workforce shrinkage—displaces all members. All members must cope first with the threatened loss of their organizations and roles, then with the actual loss and displacement that follows, and finally with the transition to new roles. Unemployment is one outcome of organizational death; less widely recognized individual transitions include retirement, transfer, changing to a job in another organization, and going back to school.

The death of an organization causes the loss of an important network of mutual obligations. It destroys a major social arena in which members have spent much of their time. Because of this loss, a closing is emotionally charged; it causes mourning, anger, depression, sorrow, and fear of the unknown, the future, and the ambiguous present.

The events leading to the death of an organization and the demise itself also place strong demands on members to make sense of an unfamiliar stream of events. In the process of disbanding an organization the cognitive structures and strategies individuals used to make sense of their organization and their roles within it become obsolete.

The threatened and actual loss of shared organizational arenas and subsequent role transitions require members to modify how they feel and think about themselves, their jobs, and their co-workers. There are many ways in which displaced members can respond to such affective and cognitive demands. They can seek individual counseling. They can rely on family and friends for emotional and financial support. Or they can withdraw. Most displaced members use some combination of these coping strategies, each of which is useful for partially meeting the demands placed upon them.

We propose that parting ceremonies offer displaced members additional help in coping with the affective and cognitive demands of organizational death. Parting ceremonies are unique coping mechanisms because many, sometimes all, displaced members can share the antidote at the same time and in the same place. Moreover, the members are at once providers and receivers of support and coping strategies.

Attributes of Parting Ceremonies

Formal parting ceremonies occurred in six of the eight dying organizations we studied. There was much variation in parting ceremonies across organizational deaths (see Table 2). The names varied; examples include the wake, the last supper, the last hurrah, and the final party. Employees paid for some events, unions for others, and management for still others. Further, as evidence from the center for the disabled suggests, parting ceremonies held for the same organizational death varied.

The ceremonies emerged in various ways. In the case of the center for the disabled, a huge picnic on the last day was suggested by a group of employees, but sponsored by management. In contrast, the local union sponsored the two major parties in the large auto plant closing. The first was with the permission of the national union, but the second party arose as an act of defiance. The national union had ordered the head of the local union to turn the money in the local treasury over to them. The local union leader said "to hell with that" and organized a bash for several hundred people.

Finally, some parting ceremonies arose despite management's resistance. The acting president of the hospital was puzzled by the emergence of a party:

> I didn't expect any parties. In fact, they asked me if the hospital would pay for and sponsor a party for the employees. . . . I said I didn't think it was appropriate to have the hospital put on a party like we would for Hospital Week or something else to celebrate. It just didn't seem like we ought to be celebrating this to me . . . so they went ahead and generated their own parties.

Despite such variation, we were intrigued by the following 11 constant elements that occurred across these parties. An example is provided for each element.

1. Participants exchange names and addresses.

   A list of former and current employees' names, addresses, and phone numbers was compiled prior to the picnic at the center for the disabled and then distributed at the picnic.

2. Participants promise to keep in touch.

   At the hospital, a nurse reported that plans were set for a yearly reunion of all former employees.

3. Participants consume food and alcoholic beverages.

   The "last hurrah" for the large auto plant was characterized by eating, drinking, and dancing "all day and all night," in the words of one informant.

4. Participants express sadness.

   "It was sad," a member of the academic unit said; "somber, sad I guess," said a former auto plant worker.

5. Participants express anger.

   The head nurse at the hospital reported: "There was a fair amount of animosity . . . more than I expected."

TABLE 2
Names and Attributes of Eleven Major Parting Ceremonies in Six Organizational Deaths

Organization	Name (if any)	Primary Attributes
Hospital	The last supper	A catered affair. Held six days before the closing. Sponsored by a physician with 20 years tenure. "Downhome country style food." All the employees were invited, even those laid off two months earlier. Cost about two thousand dollars.
	A big party at a restaurant (no name)	On last day. Sponsored by employees. Employees paid out of own pockets, but president sent money to buy each employee a drink. Much bitterness by the unemployed towards those employees absorbed by parent hospital. All employees invited. President unable to attend.
Large auto plant	Christmas party	Sponsored by union. Held at local fairgrounds. Attended by 900 people. Aim of party was to say good bye to those transferring to southern plant. Very sad occasion, not bitter.
	The last hurrah	Sponsored by union for those not going to southern plant. Union leader organized it to deplete the local union's treasury, which the national union demanded be turned over to them. Management invited. Had a bitter, defiant tone. Had food, drink, and dancing. 300-400 attendees. Lasted all day and well into the night.
Large retail store	Employee party (no name)	Got together to say good bye at a restaurant. Sad but not bitter. The store manager sent his best wishes but did not attend because he feared confrontations with employees.
	Management party (no name)	Got together to say good bye to the store. A lot of storytelling and a few teary moments. All managers knew that they would still have jobs with the parent company but there was some bitterness because they felt they could have made the store lucrative.

TABLE 2 (Continued)

Organization	Name (if any)	Primary Attributes
Center for the disabled	Recognition day	Attended by 560. Certificates of appreciation given to friends of the center. Held outside with food and drink. Funded by Center.
	The picnic	Big, formal party held on grounds of the facility. Very sad. Held after most lay-offs occurred. About 200 attended.
	The final lunch	Held at Elks Club for remaining 20 employees. Sad, but all glad the wind-down was over.
Academic unit	The wake	A final party for unit members and friends of unit. Sponsored by department. Purpose was to remember "vigor and charm of the departed spirit." Letters of support from around the country were read. One member would not attend because he didn't accept the demise.
Research organization	The final party	Held at the facility. About 100 current and former members and friends of the organization attended. Rock and roll from 60s blared. Much emotion. Much anger. People singing "The party's over, it's time to call it a day!" The president given a plaque.

6. Former members who have left are invited and attend.[2]

   An informant who attended the wake held for the academic unit reported: "Well it was nice for one reason. It was like a reunion because there were a lot of people who had been members of the organization who had come whom we hadn't seen for some years. So it was nice and felt like a high school reunion."

7. Participants discuss their own future and that of others.

   The head of the center for the disabled made the following observation about one of their parting ceremonies: "There was a lot of talk about what are you going to do next. A lot of people were asking me what I was going to do, for instance. I had headed the facility. They had some psychological investment, if you will, in my future. I guess they felt if I could find a job maybe they would too or something."

8. Participants tell one another "I guess this means it's really over."

   One informant remarked, "I've never seen a guy come back from the grave after his funeral."

9. Participants discuss the causes of the organization's demise.

   While one parting ceremony was being planned, members of the soon-to-be defunct organization designed and produced a t-shirt with an inscription that blamed the parent organization for the death. This t-shirt was the subject of much discussion at the ceremony.

10. Participants tell stories about the "old place."

    The members of the academic unit told stories about the days when the unit was rated as the best in the country, and people who could not attend the wake were asked to "please consider sending a message perhaps containing an anecdote you remember with pleasure."

11. Participants take photographs.

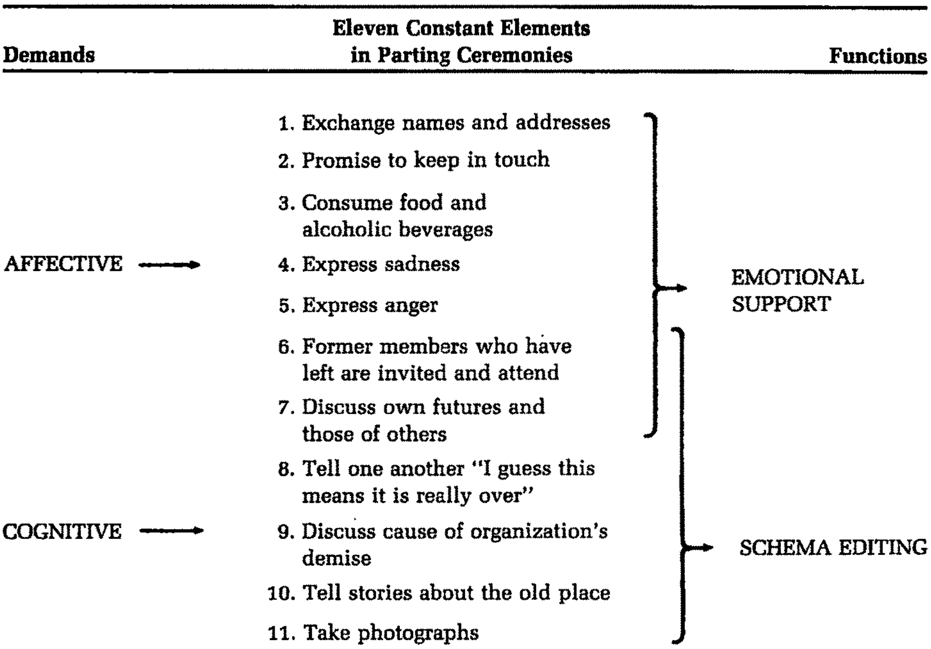
    One administrator reported: "We got a room down at the Elks Club and had a lunch, took a few pictures. There was one point where, I didn't get all 20, but I got 17 or 18. I set up my camera with the timer and tried to take a picture of the whole crew."

We propose that these 11 constant elements do not arise by coincidence. They arise in response to the demands that organizational death places on members. As Moore and Myerhoff noted, " collective ritual can be seen as an especially dramatic attempt to bring some particular part of life firmly and definitely into orderly control" (1977: 3). These common elements appear to serve functions that parallel the affective and cognitive demands faced by displaced members. They reflect settings that (1) provide emotional support, and (2) facilitate editing of schemata. Figure 1 summarizes the theoretical relationships among the affective and cognitive demands, the 11 constant elements that arise in parting ceremonies, and the two functions. As Figure 1 reveals, some constant elements may serve both functions. For instance, former members can provide both emotional support and information for schema editing. Other elements appear to serve primarily one function; for

---

[2] It was unclear from our data whether this activity occurred in the large retail store's parting ceremony.

FIGURE 1
The Functions of Parting Ceremonies in Dying Organizations: A Conceptual Framework


Demands
Eleven Constant Elements in Parting Ceremonies
Functions
1. Exchange names and addresses
2. Promise to keep in touch
3. Consume food and alcoholic beverages
AFFECTIVE
4. Express sadness
5. Express anger
EMOTIONAL SUPPORT
6. Former members who have left are invited and attend
7. Discuss own futures and those of others
8. Tell one another "I guess this means it is really over"
COGNITIVE
9. Discuss cause of organization's demise
SCHEMA EDITING
10. Tell stories about the old place
11. Take photographs

instance, taking snapshots facilitates only schema modifications. We explore the relationship between the constant elements and functions below.

Emotional Support

An administrator described the response of employees of the hospital to the closing announcement thus:

> "A lot of people cried. An awful lot of people cried, more than I expected actually."

The head nurse of the dying hospital said:

> "I made sure that I was visible to people. And if they wanted to scream at me, they could scream at me. I mean it hurts your feelings and everything. Even though you say you're not going to take it personally, you do. But I think people needed to vent."

Organizational death packs an emotional wallop. Survey data from this research support the notion that the deaths of organizations are emotionally distressing. We asked informants who had been members of the eight defunct organizations to compare the closing experience with five increasingly threatening life experiences: (1) receiving a parking ticket, (2) an argument with a friend or relative, (3) a serious illness, (4) a divorce, and (5) the death of a spouse.[3] The mean response (3.39 on a 5-point scale) indicated that the 23 former members felt that the closing experience was somewhat worse than a serious illness. Responses of 9 of the 23 informants (39%) highlight the individual distress generated by these organizational transitions; they reported that the closing experience was equivalent to a divorce or the death of a spouse.

In qualitative terms, informants spoke of "missing the old place" and how much they had "loved" their organization. The following excerpt from the *Wall Street Journal* further illustrates the sense of loss that can accompany organizational death:

> A psychologist asked Jerald H. Maxwell, former head of the defunct Medical General Inc., "If all your family was killed, how would you feel?" He replied, "Bad but losing my company hurts worse" (Ingrassia, 1982: 20).

Because the emotional responses evoked by a closing are negative and uncomfortable, members are motivated to cope with them. Emotional support can functionally counter these affective demands. Our concept of emotional support follows from House's (1981) concept of the same name and Kahn and Antonucci's (1980) concept of affective support. Emotional support includes people's expressions and subsequent perceptions of esteem, trust, liking, respect, admiration, and love for each other. It also includes their simply listening to each other.

Those who attend parting ceremonies exchange emotional support in both explicit and implicit ways. At least 7 of the 11 constant elements serve the function of emotional support, as Figure 1 shows. Exchanging addresses and promising to keep in touch suggest, albeit implicitly, that people like one another.

These acts also help members break current social bonds in a way that is graceful and socially acceptable. They exchange names and addresses and promise to keep in touch because these symbolic acts enable them to leave their old roles without embarrassment. In Goffman's (1959) terminology, it enables them to save face. It is much easier to walk away from friends with a promise to keep in touch. In contrast, people rarely say, "I'll never see you again," because that would make parting awkward and embarrassing. The administrator of the center for the disabled noted:

> it [the party] was very much like when people move away and they tell each other that they are really not saying goodbye. They promise to come back and visit but it doesn't happen. It is a way of making a transition.

Furthermore, work roles entail more than just commitment to jobs; they also require commitment to social systems. People are reluctant to give up both jobs and meaningful relationships. To ease the stress of displacement, individuals need to feel that loss of their organizational roles will not sever the overarching relationships between members. In their review of the literature on adult socialization, Mortimer and Simmons (1978) found evidence that "persons with meaningful attachments in several life areas may be protected from a sense of overwhelming loss when disengagement from any single role becomes necessary" (1978: 426). Huntington and Metcalf (1979) discussed how the continuity of life for those not dead is "focused upon and enhanced during the rituals surrounding a death" (1979: 93). Similarly, parting ceremonies provide settings in which participants can reaffirm their commitment to each other beyond the bounds imposed by an organization. This is also one of the reasons that participants discuss one anothers' futures.

---

[3] This question was part of a self-administered questionnaire used in the larger study of organizational death (Sutton, 1984). Except for using this "life events ladder" (Cobb & Kasl, 1977: B13-15), the present analysis of parting ceremonies relies exclusively on qualitative data gleaned from informants and from records data. Only informants who had been members of the eight dying organizations were asked this question; of the 30 members, 23 returned the questionnaire. Sutton (1984) contains a complete description of instruments used.

Sharing food and drink can serve as a symbol for social relationships (Gusfield & Michalowicz, 1984); thus, it is another way in which people facing adversity can reaffirm a sense of community, or "family," in the words of some informants. Parting ceremonies in this research usually included meals, not just snacks. People usually view meals as special times that should be shared with family or friends, and hence, they symbolize closer relationships than mere membership in the same organization implies.

Listening to each other express sadness and anger is one way in which members of defunct organizations exchange emotional support in these ceremonies. As Trice and Beyer (1984) pointed out, certain rites "permit expressions of emotion—including unacknowledged conflict—that are not allowed on other occasions. Letting off steam doesn't count in the same way as it would on other occasions" (1984: 663). Parting ceremonies also provide opportunities for others to respond to these sanctioned expressions with empathy and understanding.

The presence of alcoholic beverages may serve to further sanction the expression of emotions that are unacceptable in other contexts. Perhaps individuals expressed some of the anger and sadness because of the physiological influence of alcohol on inhibitions. A number of authors, however, have argued that the symbolic effects of alcohol on expression are more powerful than the physiological effects (Cavan, 1966; McAndrews & Edgerton, 1969, Van Maanen, 1984). Drinking can be viewed as a socially constructed ritual that permits certain behaviors at certain times, including the expression of otherwise unacceptable emotions. In addition to loosening people's inhibitions, wine, beer, and liquor may serve as common symbols that "can minimize for a ceremonial moment their disconnections and conflicts in a crowd" (Moore & Myerhoff, 1977: 6).

The loss of a job has been compared before to the loss of a loved one (Greenhalgh, 1979). Experiencing feelings similar to those evoked by the loss of a friend or relative, people feel both angry and depressed about losing common organizational arenas. Parting ceremonies provide socially accepted settings in which members can grieve together for the loss of their organizations. Most informants noted the analogy between parting ceremonies and funerals. Names such as the last supper and the wake also reflect this analogy.

As does the coming together of people bound by a common friend who has died, a parting ceremony enables people to say goodbye to one another as one large body. The process of mourning exhibited in the parting ceremonies we studied may also function to help "detach the survivor's memories and hopes from the dead" (Freud, 1952: 65). This sharing of a common loss and the support conveyed during this process is accentuated by former members' attending and by members' expressing concern for one another when future roles are discussed.

Finally, parting ceremonies may promote constructive social comparison. A primary reason that parting ceremonies may arise during organizational death is that people appear to seek the company of similar others when they are under stress. In a series of classic experiments, Schacter (1959) gave subjects expecting to receive electric shocks a choice between waiting alone or with others. People who anticipated receiving a painful electric shock were more likely to request the company of others than those who anticipated receiving a mild shock. He interpreted these results as evidence that people facing high stress prefer the company of similar others so that they can evaluate and adjust their responses to the threat. Most people have little experience in coping with organizational death. By interacting with others who are also trying to cope with the demise, members can assess the appropriateness of their emotional and expressive responses.

Schema Editing

It appears that parting ceremonies also provide settings that facilitate the editing of displaced members' schemata. A schema is an internal cognitive structure for encoding and representing incoming information with some efficiency (Markus, 1977). Schemata are essential to humans because they serve as initial frames of reference for behaving and perceiving (Weick, 1979).

According to Neisser's (1976) perceptual cycle, individuals' schemata act to direct perceptual explorations of their current environments. Individuals may construe the information samples garnered in such explorations as congruent or incongruent with their schemata. The perception of incongruence can lead to schema modifications that in turn lead to somewhat different perceptual explorations.

Parting ceremonies create settings that facilitate such perceptual explorations. We assert that parting ceremonies act as settings in which displaced members can collectively compare, evaluate, and modify their existing schemata. We propose that at least six of the behavioral transactions that occur among those who attend parting ceremonies facilitate the editing of three interrelated schemata. The three are social system, self-, and event schemata. Although these schemata are not mutually exclusive, it is useful to describe them separately for analytical purposes.

Social system schemata. A social system schema is a displaced member's cognitive map of the history, attributes, and fate of an organization. Parting ceremonies provide many opportunities for people to edit their social system schemata.

In all parting ceremonies we studied, people spent much time telling stories about their organizations and its members. This storytelling did more than make people smile. It also helped them to remember, compare mental histories, and edit and encode their own stories. Parting ceremonies can function as "summaries," a term borrowed from Albert and Kessler:

> The summary is defined as an abbreviated statement of the history of a social encounter. More specifically, it is a "selective repetition" of aspects derived from the "history" of the social encounter. This repetition is presented in a "condensed" and often "symbolic" form .... The summary helps the departing participants remember the encounter. A summary processes historic information in such a way that it allows the participants to symbolically take the encounter away with them as they leave (1976: 154–155).

Albert and Kessler also noted that because summaries are selective, they make possible abbreviated histories of encounters. Further, the omission of certain events due to abbreviation allows for their devaluation.

Slote (1969) provided an anecdote from a parting ceremony that conveys how telling stories can help people edit their mental histories:

> Everyone told his own story, the funny things, the angry things, stories about guys who'd left fifteen to twenty years ago, and what happened to old so-and-so who worked in S and R and used to pitch horseshoes behind the boiler room, incidents out of the past came tumbling out, fights, ball games. Perry DeMoss came in and had a drink and pretty soon he was claiming he was the best second-baseman that Baker ever had, and on it came: old Christmas parties; old secretaries; rush orders from Chrysler that had kept them hopping week-end nights; the big snowstorm in 1952, and the time Ramsey and Paddleford backed a truck into forty drums and the drums went rolling every which way and Ramsey laughed and laughed; and the time the trucks got stuck in the yard during an ice storm and the drums started rolling off; Steve Toucek's fight with what's-his-name .... And on it went like that, stories that were partly true, maybe even mostly true. I knew then, watching the men drink and listening to each man tell his tale, that each of them had his own Baker, each had his own plant, and each had his own memories to last him the rest of this life. It had been a good place to work—for all of us (1969: 319).

Mental imagery was not the only way that people recorded the histories of the social system. Amateur photographers were ubiquitous in the parting ceremonies we studied. The set of people included in each photograph appeared to be intentional. Usually at least one photograph included the entire group. And many subgroup photos were taken, that often captured functional, departmental, and informal structures of defunct organizations. As anyone knows who has spent an evening perusing family photo albums, photography is an exercise in preserving and editing history.

In parting ceremonies, as in all phases of organizational death, members spend much time discussing why an organization died and who was to blame, in order to make sense of an important event in their lives. This is an essential component of the cognitive summaries that displaced members leave with. Humans have a strong need to know and understand the causes of events, and research suggests (Nisbett & Ross, 1980) that humans seek to develop causal hypotheses about events from even the most meager clues, especially when the events are important to them. Parting ceremonies offer perhaps final opportunities to reaffirm or adjust hypotheses about the causes of an organizational death. For example, at the wake held for the academic unit, members reaffirmed the "evil" role of the parent organization. And at the last supper held by hospital employees, different subgroups developed and reaffirmed their different hypotheses about the demise.

At least two of the organizational leaders in our research understood well the human predilection towards placing blame on individuals—particularly managers (Pfeffer, 1977)—rather than on events (Ross, 1977). The manager of the large retail store and the acting president of the hospital both chose not to attend parting ceremonies because they feared confrontations with displaced workers who blamed them for the deaths.

Finally, as Turner (1969) asserted, ceremonies do not always serve to reinforce static sets of relationships but can also mark shifts in those relationships. As Moore and Myerhoff observed: "Since ritual is a good form for conveying a message as if it were unquestionable, it often is used to communicate those very things which are most in doubt" (1977: 24). In all parting ceremonies we studied, people told one another "I guess this means it is really over." A parting ceremony provides unambiguous and vivid evidence of an organization's demise. This helps individuals avoid unhealthy denial of a closing, which is common in organizational death (Slote, 1969; Taber et al., 1979).

Self-schemata. Self-schemata are cognitive generalizations about the self, derived from past experience, that organize and guide the processing of self-related information contained in an individual's social experiences (Markus, 1977). The concept of self is a function of the qualities and outcomes of an individual's experience. Telling stories about their experiences and noting the reactions of others allows displaced members to edit their histories as members of the dying organization. Displaced members may hear others tell stories about themselves; these stories may lead to affirmation or modification of self-histories. This storytelling also serves to adjust or reaffirm displaced members' thoughts about qualities such as their ability to tell jokes, be creative, work hard, be lazy, or help others.

It is important to self-schemata that the stories focus on individual actions. Organizational stories often contain morals (Bower, 1976; Martin, 1982). Stories that emerge in parting ceremonies seem to stress the moral that each individual played an important role in shaping the experiences that all shared.

Parting ceremonies also provide opportunities for displaced members to reaffirm or question beliefs about their responsibility for their organization's demise. For example, a ceremony provided an opportunity for a manager of the retail store to remind others, and thus himself, that the organization died despite his fine work to reduce the store's losses.

P3955

The events of parting ceremonies allow the displaced members to validate and edit their self-schemata through reification of their previous organizational experiences. Moreover, such ceremonies serve as settings in which they can begin to interpret the implications of the closing event for self. Slote (1969) designated displaced members as either winners or losers. Parting ceremonies allow individuals to gather data that help them apply such labels to themselves. For example, one hospital administrator who received a promotion from the parent hospital reported being taunted by many of the other displaced members who had no jobs: "You've got a job to go back to, you know the hospital is gonna pick you up. Why can't the hospital assume all of us?" She responded to these remarks by telling us "I started to feel guilty about being employed."

In literary terms, parting ceremonies help displaced members complete chapters in their mental autobiographies—chapters about their tenures in the defunct organizations. These resemble the abbreviated summaries of social encounters that Albert and Kessler (1976) described.

Parting ceremonies also facilitate planning for the next chapters of these autobiographies. By talking with people who have already successfully made transitions from organizational membership to new roles, displaced members can cope with the poor self-images that may result from association with failure and can begin to plan and picture desired selves (Wicklund, 1975).

Event Schemata. Event schemata, or scripts, are like cartoon strips that summarize "basic actions that can be executed in a range of possible manners and contexts" (Nisbett & Ross, 1980). Event schemata help people decide what to do next in given situations. Upcoming or current role transitions evoked by organizational death may include retiring, starting a new job in a parent organization, starting a new job in another organization, looking for a new job, going to school, or collecting unemployment benefits. If, as Abelson argued, a script is "a coherent sequence of events expected by the individual" (1976: 34), then most people have incomplete, inappropriate, obsolete, or nonexistent event schemata to guide their behavior through these transitions. Extensive research suggests that people experience distress when they are unable to predict future events that are important to them (e.g., Seligman, 1975; Sutton & Kahn, 1986).

Parting ceremonies create settings in which displaced members can explore with each other their ideas for dealing with role transitions. The displaced members we studied spent much time talking to one another about the future. In a period of individual transition, networks of acquaintances can provide timely and important information about job possibilities, influential people, or even information about how to sign up for unemployment benefits. Informants also reported that parting ceremonies were used to get letters of recommendation from former and current supervisors and phone numbers of "people you should talk to."

Former members who had already made transitions to new roles were invited to attend parting ceremonies. These former members spent much time discussing their own transitions and their outcomes. They were often asked to judge the quality of coping strategies developed by members who had yet to make a transition. A typical question might be, "So what do you think about this approach to finding a job?" In short, parting ceremonies provide settings that can help displaced members anticipate the future and learn what actions are appropriate for that future.

FUNCTIONS OF PARTING CEREMONIES FOR THE MANAGEMENT OF DYING ORGANIZATIONS AND FOR PARENT ORGANIZATIONS

Our emphasis thus far has been on how parting ceremonies help individuals cope with the displacement caused by organizational death. We assert that, in addition to the benefits they offer displaced members, parting ceremonies may also serve functions for the managements of dying organizations and for their parent organizations, if such exist. The theoretical foundations of this assertion are less well developed than those for the individual functions. It is important that we describe some of these functions, however, because so little is known about the management of organizational death (Sutton, 1983).

We will focus on how management can influence parting ceremonies to enhance such functions. However, it is essential to distinguish functions from managerial actions. Managers may seek to create ceremonies that enhance outcomes they desire, but links between organizational changes and managerial intentions can be remarkably weak (March, 1981). Moreover, it may not occur to some managers that parting ceremonies can be points of intervention. As mentioned earlier, the president of the dying hospital was horrified to learn that employees were "celebrating" the demise. Nonetheless, some leaders do appear skilled at orchestrating such ceremonies. The head of the center for the disabled, for example, used Recognition Day and the picnic to help maintain employee commitment and effort. And such actions were taken with skill and forethought.

Specifically, we propose that skilled managers can use parting ceremonies in order to influence five functions: (1) member motivation, (2) information dissemination, (3) external stakeholders' acceptance, (4) impression management, and (5) guilt assuagement.

Member Motivation

Parting ceremonies, particularly those that managements plan and sponsor, may satisfy organizational and societal norms about the treatment of members. Sponsoring a parting ceremony requires resources, particularly time. Employees know that resources are usually scarce during organizational death. According to the norm of reciprocity (Blau, 1964), when management takes time to plan a ceremony, employees may put forth greater effort to complete remaining tasks. Parting ceremonies are also often used to recognize people who have been important to an organization. Since the organization's formal reward base—job security, advancement, and so forth—is often eroded, a ceremony allows management to give a collective pat on the back to lower level employees.

Perhaps more important, parting ceremonies provide management with settings in which they can smooth over or resolve conflicts with employees. Management can use the occasion to break down the barriers that separate them from employees and promote the image of a shared fate. It is important that management be able to come across as less villainous and less well-off than they appear to be to employees who may feel betrayed or used. Restoring good will is a crucial step in insuring that employees will continue to do their jobs and will not engage in counterproductive behaviors such as sabotage or theft.

Information Dissemination

Parting ceremonies facilitate the sharing of information. Since usually most members are present, management can send the same message to all through memos or announcements. At the picnic held at the center for the disabled, for example, a handout was passed out that described how to get unemployment benefits. A more subtle way in which parting ceremonies facilitate the dissemination of information was mentioned earlier. A ceremony serves as a symbolic way of saying a closing is a reality.

External Stakeholder Acceptance

Parting ceremonies also make explicit to other stakeholders that a closing is truly going to occur. Stakeholders may be invited to parting ceremonies. For example, when parents of residents attended Recognition Day at the center for the disabled, it forced them to face hard evidence that the center was going to close. Students from the academic unit attended the wake, which had a similar effect. Further, when members of other stakeholding groups get indirect word of a parting ceremony, it helps them understand that a closing is actually going to occur. For example, when suppliers of the large auto plant heard that 900 employees had held a party to say goodbye to one another, it affirmed that the plant was going to close. Thus, parting ceremonies can have a ripple effect that enhances sense-making among other stakeholders.

Impression Management

When dying organizations are subunits of larger organizations, management can use parting ceremonies to create favorable impressions among external stakeholders. Providing employees with a nice picnic and giving them certificates of merit, as the center for the disabled did, signaled other stakeholders that the center was breaking links in ways that were appropriate to organizational and societal norms. Breaking links in a normative way, or in ways that are more generous than existing norms suggest, is important for parent organizations that depend on support from customers or clients for survival. In this sense, parting ceremonies become stages on which management can further their interests by presenting "administrative dramas" that convey messages favorable to them (Meyer, 1982: 533).

Guilt Assuagement

Many of the managers with whom we spoke felt guilty about the closing and its negative effects on displaced members. The head of the center for the disabled was very upset because almost 70 percent of those whom he had helped lay off were collecting unemployment insurance. He also felt guilty because the center had not taken enough steps to recognize individual employees. Similarly, Slote (1969) reported that the guilt of closing the Baker plant led Frank Robertson, the last plant manager, to experience severe physical distress.

Impression management refers to influencing the evaluations of outsiders, but assuaging guilt refers to the internal reactions of management to parting ceremonies. Sponsoring formal ceremonies in which they can buy drinks, buy food, offer thanks, or even offer letters of recommendation for displaced employees can help managers reestablish a sense of equity and therefore feel less guilt.

DISCUSSION

Parting ceremonies that arise during organizational death have been described as coping mechanisms for displaced members and as points of intervention that leaders can use to manage organizational death. We have proposed that parting ceremonies help members cope with the affective demands of organizational death by creating settings in which members can provide each other with emotional support. We have also proposed that these ceremonies help members cope with the cognitive demands of organizational death because they facilitate the editing of social system, self-, and event schemata. Finally, we have described five functions that parting ceremonies can serve for managements of dying organizations and their parents.

These propositions require qualification, expansion, and testing. Our aim was to build a data-based theory. Glaser and Strauss (1967) asserted that theories grounded in empirical evidence should be superior to those grounded only in the imagination. Despite its empirical grounding, the theory offered here is not complete. We contend, however, that it provides a useful point of departure for subsequent research and theory-building efforts.

Perhaps the theory's most evident limitation is the lack of attention it gives to negative or dysfunctional aspects of parting ceremonies. We implied earlier that some people benefit more from parting ceremonies than others. Although placing the blame for a closing on others may complete mental pictures for some, the impact on those blamed is likely to be negative. In parting ceremonies, as in organizational death itself, there appear to be winners and losers.

Why did this study lead us to such a strong emphasis on positive consequences? This emphasis may reflect a systematic effort by decision makers to justify their actions rather than a ubiquitous quality of parting ceremonies. Yet this potential source of distortion is attenuated in this study because decision makers often had little stake in the gatherings; people outside of the managerial ranks instigated at least 6 of the 11 ceremonies. Moreover, nonmanagers like union leaders and secretaries also emphasized the positive consequences of parting ceremonies.

Borrowing Gestalt terminology, we could say that the positive consequences of parting ceremonies stand out against the negative background created by the death of an organization. The negative consequences of parting ceremonies, on the other hand, may blend into this background and not register as significant. To more fully capture the complexity of parting ceremonies, future research and theory-building efforts need to focus on the interaction between the negative and positive consequences such ceremonies have for individuals and managements of dying organizations.

Another limiting feature of the theory offered here is its neglect of variations among parting ceremonies. As Udy (1965) observed, theories about organizations can be divided into those emphasizing variables and those emphasizing constants. Our focus on constants facilitated building a theory that captures the basic structure of parting ceremonies and provides initial support for the pervasiveness of these phenomena. It is difficult to examine features that differ across settings without an idea of how the settings are similar. Thus, the theory offered here may help guide research on the differences among parting ceremonies.

Subsequent theory-building might also address differences between organizational deaths in which parting ceremonies occur and organizational deaths in which they do not occur. We observed no parting ceremonies in two of the eight organizational deaths we studied—the small auto plant and small sporting goods store. The qualifier "small" in the description of these two organizations is telling. Perhaps efficiency is one reason that formal parting ceremonies arise—they provide settings in which most or all members of dying organizations can see one another for the last time. In the small retail store, the three members saw one another almost every minute of the working day and could handle parting informally. Perhaps formal parting ceremonies only arise when there is a need to supplement informal meetings. Similarly, in their discussion of ending social encounters, Albert and Kessler (1976) asserted that group rituals arise when individual summaries are too difficult to develop and when a group ritual is more efficient.

The absence of a parting ceremony in the small auto plant may have been due to smallness; however, other qualities of this organizational death may have eliminated the forces that give rise to such ceremonies. Parting ceremonies seem to arise to address the shared needs of members who have been or are about to be displaced. The small auto plant's death was gradual, and no one ever acknowledged that the organization was dying. This gradual process may have eliminated the feeling of shared fate that gives rise to parting ceremonies. In addition, the owner of the plant engaged in denial about the impending fact of organizational death. Perhaps a parting ceremony would have helped him and other members accept the demise of this organization.

It is also possible that no parting ceremonies emerged in these two organizations because of some feature inherent in private sector organizations that are independent of larger organizations. As Table 1 reveals, the small sporting goods store and the small auto plant were the only two organizations studied that fit this category. Perhaps parting ceremonies failed to arise because there was no need to satisfy the norms of a parent organization or a public agency. The present study does not allow us to determine which if any of these possibilities explain the absence of parting ceremonies. Future research should focus on the emergence of parting ceremonies and the differences in circumstances and outcomes between organizational deaths that include parting ceremonies and those that do not.

More broadly, this study suggests that organizational scholars should ask how formal occasions, as a class of ceremonies, can enhance organizational outcomes. Some anecdotes suggest that effective companies use parties to socialize and gain loyalty from members. Tandem's Friday "beer busts" are one example (Deal & Kennedy, 1982).

We need more evidence about the link between ceremonies and organizational outcomes. Some companies spend huge sums on ceremonies. During the Christmas of 1984, for example, Advanced Micro Devices spent $700,000 on a single party for 10,000 guests, and Apple Computers spent over $110,000 for 19 parties (Bellew, 1984). Is such money wasted? Or, as executives from these companies claim, are such extravaganzas essential for attracting the best talent, maintaining commitment, and keeping the best employees from jumping ship?

The research described here and the theory that follows from it suggest that formal ceremonies play an important role in organizational death and in other domains of organizational life. We have argued that parting ceremonies have benefits for displaced members, for management, and for external stakeholders. This perspective suggests that subsequent theory-building and testing should explore the opportunities and hazards that formal ceremonies hold for the well-being of organizations and their members. Only in this way can we begin to assess the psychological and economic value of Friday beer busts, Christmas parties, sales conventions, and annual company picnics.

REFERENCES

Albert, S., & Kessler, S. 1976. Processes for ending social encounters: The conceptual archaeology of a temporal place. ***Journal for the Theory of Social Behavior***, 6: 147–170.

Abelson, R. P. 1976. Script processing in attitude formation and decision making. In J. S. Carroll & J. W. Payne (Eds.), ***Cognition and social behavior***: 33–45. Hillsdale, N. J.: Erlbaum.

Bellew, P. A. 1984. The office party is one thing at which Silicon Valley excels. ***Wall Street Journal***, 65(49): 1.

Blau, P. M. 1964. ***Exchange and power in social life***. New York: John Wiley & Sons.

Bower, G. H. 1976. Experiments on story understanding and recall. ***Quarterly Journal of Experimental Psychology***, 28: 511–534.

Cameron, K. S., & Whetten, D. A. 1982. ***Organizational effectiveness: A comparison of multiple models***. New York: Academic Press.

Cavan, S. 1966. ***Liquor license: An ethnography of bar behavior***. Chicago: Aldine.

Cobb, S., & Kasl, S. V. 1977. ***Termination: The consequences of job loss***. HEW Publication No. (NIOSH) 77–224. Washington, D.C.: U.S. Department of Health, Education, and Welfare.

Davis, M. 1971. That's interesting! Towards a phenomenology of sociology and a sociology of phenomenology. ***Philosophy of the Social Sciences***, 4: 309–344.

Deal, T. E., & Kennedy, A. A. 1982. ***Corporate cultures***. Reading, Mass.: Addison-Wesley.

Freud, S. 1952. ***Totem and taboo***, New York: Norton.

Glaser, B., & Strauss, A. 1967. ***The discovery of grounded theory: Strategies for qualitative research***. London: Wiedenfeld and Nicholson.

Goffman, E. 1959. ***The presentation of self in everyday life***. New York: Doubleday.

Goodman, P. S., Pennings, J. M., & Associates. 1977. ***New perspectives on organizational effectiveness***. San Francisco: Jossey-Bass.

Greenhalgh, L. 1979. ***Job security and the disinvolvement syndrome: An exploration of patterns of worker behavior under conditions of anticipated grieving over job loss***. Unpublished doctoral dissertation, Cornell University, Ithaca, N. Y.

Gusfield, J. R., & Michalowicz, J. 1984. Secular symbolism: Studies of ritual, ceremony, and the symbolic order in modern life. ***Annual Review of Sociology***, 10: 417–435.

Hannan, M. T., & Freeman, J. 1984. Structural inertia and organizational change. ***American Sociological Review***, 49: 149–164.

Hardy, C. 1982. ***Organizational closure: A political perspective***. Unpublished doctoral dissertation, Warwick University, Coventry, England.

House, J. S. 1981. ***Work stress and social support***. Reading, Mass.: Addison-Wesley.

Huntington, R., & Metcalf, P. 1979. ***Celebrations of death***. New York: Cambridge University Press.

Ingrassia, L. 1982. Aftermath of failure: The collapse of a man, the agony of a family. ***Wall Street Journal***, 62(104): 1–20.

Kahn, R. L., & Antonucci, T. 1980. Convoys over the life course: Attachment, roles, and social support. In P. B. Bates & O. Bim (Eds.), ***Life span development and behavior***, vol. 3: 254–283. Boston: Lexington Press.

Katz, D., & Kahn, R. L. 1978. ***The social psychology of organizations*** (2nd ed.). New York: John Wiley & Sons.

Lawrence, P. R., & Lorsch, J. W. 1967. ***Organization and environment***. Boston: Harvard Business School, Division of Research.

Lawrence, P. R. 1981. The Harvard organization and environment research program. In A. H. Van de Ven & W. F. Joyce (Eds.), ***Perspectives on organization design and behavior***: 311–337. New York: John Wiley & Sons.

March, J. G. 1981. Footnotes to organizational change. ***Administrative Science Quarterly***, 26: 563–577.

Markus, H. 1977. Self-schemata and processing information about self. ***Journal of Personality and Social Psychology***, 35: 63–77.

Martin, J. 1982. Stories and scripts in organizational settings. In A. Hastorf & A. Isen (Eds.), ***Cognitive social psychology***: 255–305. New York: Elsevier-North Holland, Inc..

McAndrews, C., & Edgerton, R. B. 1969. ***Drunken comportment: A social explanation***. Chicago: Aldine.

McKelvey, B. 1982. ***Organizational systematics: Taxonomy, evolution, classification***. Berkeley: University of California Press.

Meyer, A. D. 1982. Adapting to environmental jolts. ***Administrative Science Quarterly***, 27: 515–537.

Miles, M. B. 1979. Qualitative data as an attractive nuisance. ***Administrative Science Quarterly***, 24: 590–601.

Miller, G. 1985. Work, ritual structures, and the legitimization of alternative communities. ***Work and Occupations***, 12: 3–22.

Miller, J. G. 1978. ***Living systems***. New York: McGraw-Hill Book Co.

Mintzberg, H. 1979. An emerging strategy of "direct" research. ***Administrative Science Quarterly***, 24: 580–589.

Moore, S. F., & Myerhoff, B. G. 1977. Secular ritual: Forms and meanings. In S. F. Moore & B. G. Myerhoff (Eds.), ***Secular ritual***: 3–25. Amsterdam: Van Gorcum.

Morgan, G., Frost, P. J., & Pondy, L. R. 1983. Organizational symbolism. In L. R. Pondy, P. J. Frost, G. Morgan, & T. C. Dandridge (Eds.), ***Organizational symbolism***: 3–38. Greenwich, Conn.: JAI Press.

Mortimer, J. T., & Simmons, R. G. 1978. Adult socialization. In R. Turner, J. Coleman, & R. C. Fox (Eds.), ***Annual review of sociology***, vol. 4: 421–454. Palo Alto, Calif.: Annual Reviews, Inc..

Nance, J. J. 1984. ***Splash of colors***. New York: Morrow.

Neisser, U. 1976. ***Cognition and reality: Principles and implications of cognitive psychology***. San Francisco: Freeman.

Nisbett, R. & Ross, L. 1980. ***Human inference: Strategies and shortcomings of social judgment***. Englewood Cliffs, N.J.: Prentice-Hall.

Pfeffer, J. 1977. The ambiguity of leadership. ***Academy of Management Review***, 2: 104–112.

Ross, L. 1977. The intuitive psychologist and his shortcomings. In L. Berkowitz (Ed.), ***Advances in experimental social psychology***, vol. 10. New York: Academic Press.

Schacter, S. 1959. ***The psychology of affiliation***. Stanford, Calif.: Stanford University Press.

Seligman, M. E. P. 1975. ***Helplessness***. San Francisco, Calif.: W. H. Freeman & Co.

Slote, A. 1969. ***Termination: The closing at Baker plant***. Indianapolis, Ind.: Bobbs-Merrill.

Starbuck, W. H., & Nystrom, P. C. 1981. Designing and understanding organizations. In P. C. Nystrom & W. H. Starbuck (Eds.), ***Handbook of organizational design***, vol. 1: ix–xxii. New York: Oxford University Press.

Sutton, R. I. 1983. Managing organizational death. ***Human Resources Management***, 22: 391–412.

Sutton, R. I. 1984. ***Organizational death***. Unpublished doctoral dissertation, University of Michigan, Ann Arbor.

Sutton, R. I., & Kahn, R. L. 1986. Predicting, understanding, and control as anecdotes to organizational stress. In J. Lorsch (Ed.), ***The handbook of organizational behavior***: forthcoming. New York: Prentice-Hall.

Taber, T. D., Walsh, J. T., & Cooke, R. A. 1979. Developing a community-based program for reducing the social impact of a plant closing. ***Journal of Applied Behavioral Science***, 15: 133–155.

Tichy, N. M. 1981. Networks in organizations. In P. C. Nystrom & W. H. Starbuck (Eds.), ***Handbook of organizational design***, Vol. 2: 225–249. New York: Oxford University Press.

Trice, H. M. 1985. Rites and ceremonials in organizational culture. In S. B. Bacharach & S. M. Mitchell (Eds.), ***Research in the sociology of organizations***, vol. 4: 221–270. Greenwich, Conn.: JAI Press.

Trice, H. M., Belasco, J., & Alutto, J. M. 1969. The role of ceremonials in organizational behavior. ***Industrial and Labor Relations Review***, 23 (October): 40–51.

Trice, H. M., & Beyer, J. M. 1984. Studying organizational cultures through rites and ceremonials. ***The Academy of Management Review***, 9: 653–669.

Turner, T. 1977. Transformation, hierarchy, and transcendence: A reformulation of Van Gennep's model of the structure of rites de passage. In S. F. Moore & B. G. Myerhoff (Eds.), ***Secular ritual***: 53–70. Amsterdam: Van Gorcum.

Turner, V. W. 1969. ***The ritual process***. Chicago: Aldine.

Udy, S. H. 1965. The comparative analysis of organizations. In J. G. March (Ed.), ***Handbook of organizations***: 678–709. Chicago: Rand McNally & Co.

Van Gennep, A. 1960. ***The rites of passage***. Chicago: University of Chicago Press. Translated by M. B. Vizedom & G. L. Caffee. Originally published as *Les rites de passage*. Paris: Emile Mourry, 1909.

Van Maanen, J. 1984. ***Power in the bottle***. Paper presented at the Symposium on the Functions of Executive Power, Case Western Reserve University, Cleveland, Ohio.

Weick, K. E. 1979. ***The social psychology of organizing*** (2nd ed.). Reading, Mass.: Addison-Wesley.

Wicklund, R. A. 1975. Objective self-awareness. In L. Berkowitz (Ed.), ***Advances in experimental social psychology***, Vol. 8: 233–275. New York: Academic Press.

Stanley G. Harris is in the Ph.D. program in organizational psychology at the University of Michigan. His current interests include organizational change, organizational culture, and the integration of the concept of emotion into the study of organizations.

Robert I. Sutton received his Ph.D. degree in organizational psychology from the University of Michigan. He is Assistant Professor in the Department of Industrial Engineering and Engineering Management at Stanford University. His research interests include organizational decline and death, occupational stress, and the expression of emotion in organizational life.

© *Academy of Management Journal*
1986, Vol. 29, No. 1, 31–50.

DUAL COMMITMENT AND LABOR-MANAGEMENT RELATIONSHIP CLIMATES

HAROLD L. ANGLE
University of Minnesota
JAMES L. PERRY
Indiana University

In 22 municipal bus companies, the extent of union members' dual commitment—to their unions and to their employing organizations—covaried with interorganizational differences in labor-management climates. Dual commitment was higher in cooperative climates than in less cooperative ones. The extent of members' participation in the union process moderated this relationship. It was relatively strong for more active participants and relatively weak for the less active. Separate organizational and union commitments also covaried with labor-management relationship climate but in a less monotonic manner than dual commitment did.

Can unionized employees develop simultaneous commitments to their unions and their employing organizations, or is there an inherent conflict of allegiance in such dual-membership situations that forces an either-or choice? This was the essential question that formed the basis for a flurry of research on dual loyalty or dual allegiance some 25 years ago. That research left the issue largely unresolved. This study once again analyzes the dual commitment concept in relation to two variables: (1) ambient labor-management relationship climate, and (2) extent of members' participation in union activities. This study is the first to use independent sources to measure labor-management relationship climates and dual commitment.

THEORETICAL FRAMEWORK

During the 1950s and early 1960s, an apparent concern that the rapid growth of the unionization might compete with employees' commitment to their employing organizations stimulated much research on dual allegiance or dual commitment (Gallagher, 1984). Early work by such scholars as Dean (1954) and Purcell (1960) indicated that most workers show allegiance to both union and management. Dean's analysis of three unionized plants incorporated an extended period of observation and interviews with union and

---

The authors wish to thank the U.S. Department of Transportation, Office of University Research, and the Institute of Transportation Studies, University of California, Irvine for their generous support of this research. A version of this paper was presented at the 1984 Annual Meeting of the Academy of Management, Boston.

management officials and rank-and-file members, followed by a questionnaire survey. In all three plants she found that those workers who were most favorable toward management also tended to be the most likely to give their union leaders high ratings. However, as Dean acknowledged, it was unfortunate that she asked a direct question about dual loyalty in only one of the plants. Purcell's landmark study, involving a single unionized meat packing plant in Chicago, was even more restricted. As Dean did with most of her data, Purcell assessed dual allegiance by tallying proportions of workers whose responses to questions tended to be favorable toward both the company and the union.

Other scholars of the period, including Barkin (1950), England (1960), and Kornhauser, Sheppard, and Mayer (1956) argued against workers' ability to display simultaneous loyalty to both systems. Barkin, for example, asserted that the notion that employees can maintain loyalty to companies and to unions concurrently "does not stand up under even a cursory examination" (1950: 64).

After about 1960, the issue of dual allegiance lay essentially unresolved and dormant for about a quarter of a century, until a new generation of researchers raised it again. In the intervening period, a strong research theme developed around the concept of organizational commitment (Porter, Steers, Mowday, & Boulian, 1974)—a concept having obvious relevance to dual allegiance. Indeed, it has been central to recent research on dual allegiance (Fukami & Larson, 1982, 1984; Gallagher, 1984; Martin, 1981; Martin, Mageneau, & Peterson, 1982).

Reminiscent of the earlier work of Dean (1954) and Purcell (1960), such recent research has tended to measure commitment to employing organizations and commitment to unions separately and has then combined these measures to show the extent of dual commitment. For instance, Gallagher (1984) tallied employees showing strong or moderate positive responses to separate union commitment and organizational commitment scales and used the proportion that fell into these classifications on both scales as an index of dual commitment. In Martin's (1981) study of a federal government facility, questionnaires that included scales measuring attitudes toward unions and toward management[1] were administered to union members and nonmembers. This research assessed dual allegiance by recording the proportion of employees whose responses were favorable toward both referents. Fukami and Larson (1982) used the correlation between employees' scores on union commitment and organizational commitment to measure dual commitment.

Commitment has been anything but a consensual construct in the general literature. It has been described in such diverse ways as: attachment to and identification with an organization (Buchanan, 1974); a binding of attitude and belief to prior behaviors (Salancik, 1977); a resignation to a status, under penalty of forfeiture of the costs of attainment of that status (Becker, 1960); and a dependency relationship in which maintenance of an individual's "internal being requires behavior that supports the social order" (Kanter, 1972: 66). Such diverse conceptual frameworks and their operational definitions—including several variations on why and how people become committed and how to measure commitment (Angle & Perry, 1983; Kiesler, 1971; Ritzer & Trice, 1969; Stevens, Beyer, & Trice, 1978)—have created a situation in which it is often difficult to make direct comparisons among commitment studies.

---

[1] These scales were adapted from Stagner, Chalmers, and Derber (1958).

Diversity in conceptualizations and measurements of commitment has had clear effects on the coherence of recent research on dual commitment. For example, although both Gallagher (1984) and Fukami and Larson (1982) measured dual commitment by combining scores on separate measures of commitment to union and organization, they measured commitment in different ways. Gallagher employed a scale based on the work of Porter and colleagues (1974), but Fukami and Larson adopted a measure patterned after the work of Ritzer and Trice (1969). These two scales rest on quite different conceptualizations of commitment and tend to be only moderately correlated (Ferris & Aranya, 1983), making direct comparison problematic. The underlying distinction between these conceptualizations regards the extent to which an affective, as opposed to a calculative, attachment is presumed. This duality resembles Etzioni's (1975) distinction between moral and calculative involvement. The instrument developed by Porter and colleagues emphasizes members' moral involvement, but Ritzer and Trice's approach appears to have a more nearly calculative involvement as its premise (Ferris & Aranya, 1983).

In calculative involvement, bonding between member and social system results from a committed individual's linking the relationship to extrinsic outcomes on the basis of costs and benefits. Becker (1960), for example, defined commitment in terms of aversion to bearing costs of forfeiture. This refers to losses of investments or side bets that would be incurred in severing a relationship. Such an instrumental orientation seems incongruent with common definitions of allegiance or loyalty, which usually connote an affective attachment or sense of devotion.

The perspective on organizational commitment taken by Porter and his colleagues (Mowday, Steers, & Porter, 1979, 1982; Porter et al., 1974) comes closer to such notions of loyalty and allegiance. This organizational behavior approach (Staw, 1977) essentially treats commitment in terms of individuals' psychological attachments to social systems. According to Porter and his colleagues, a committed employee: (1) has a strong desire to remain a member of his or her organization; (2) internalizes the values and goals of that organization; and (3) is willing to work extra hard on behalf of the organization. Thus, attachment to membership derives not from economic exchange, but from such processes as identification and internalization (Kelman, 1958)—processes more clearly consonant with loyalty or allegiance.

A great deal of published empirical research across a range of organizational and employee types has employed the measurement procedure associated with this approach to commitment (Mowday et al., 1979). Thus there is a substantial amount of normative and psychometric information regarding the measure. Furthermore, a major research effort on commitment to unions appears to have its roots in this framework (Gordon, Philpot, Burt, Thompson, & Spiller, 1980; Ladd, Gordon, Beauvais, & Morgan, 1982).

In addition to studies of organizational commitment, the period between the two waves of research on dual allegiance or dual commitment has seen at least two other streams of research develop that appear clearly relevant to the issue. One such research domain is role conflict (Kahn, Wolfe, Quinn, Snoek, & Rosenthal, 1964; Rizzo, House, & Lirtzman, 1970). A particularly relevant form of role conflict in the present context would appear to be interrole conflict, in which an individual simultaneously occupies two or more roles that may make incompatible demands; the roles of union member and organization member are an example. Lieberman's (1956) famous study of attitude shifts illustrated how group membership can affect attitudes and values and in particular how such values differ between union and management identifications. He studied individuals as they changed their primary roles, moving back and forth between being workers and either union stewards or foremen.

The other stream of research involves a distinction between cosmopolitans and locals. This concept concerns the dilemmas facing professionals whose values and loyalties to organizations and professions sometimes collide (Gouldner, 1957; Miller & Wager, 1971). In a classic study of the occupation of personnel manager, Ritzer and Trice (1969) illustrated that the issue of dual allegiance is not limited to labor-management relations, but also clearly applies to organizational versus occupational commitments.

Gouldner (1957) and other early researchers saw cosmopolitanism and localism as decidedly zero-sum. Cosmopolitanism and localism appeared to be antithetical; strong attachment to a profession precluded attachment to an organization and vice versa. However, more recent research has suggested that an either-or choice is not inherent in these orientations. Ritzer and Trice (1969) found personnel managers to be committed to both their occupation and their employing organizations "and only slightly more to their occupation" (1969: 33). Miller and Wager (1971) held that the two orientations need not be mutually exclusive when the expectations and role demands of profession and organization remain congruent—if, for example, an organization does not demand behavior contrary to professional norms. Accordingly, cosmopolitanism and localism need not be viewed as incompatible. Rather, they may coexist, contingent perhaps upon the extent to which individuals can avoid choice dilemmas or fundamental role conflicts.

These two concepts, role conflict and cosmopolitanism versus localism, are obviously relevant to the study of labor-management relations to the extent that unions and employing organizations force members to choose between incompatible values or make incompatible behavioral demands. There is substantial evidence that various union-management interfaces differ systematically with respect to how compatible or incompatible the perceived orientations of the two parties are. Scholars have long held that they can characterize labor-management relations by the relative amounts of ambient conflict or cooperation that reside in the two parties' orientations toward one another. This background condition has been called attitudinal climate (Chalmers, Chandler, McQuitty, Stagner, Wray, & Derber, 1954), models of labor management relations (Harbison & Coleman, 1951), or, perhaps most commonly, labor-management relationship patterns (Walton & McKersie, 1965). Whatever the terminology, the underlying idea has been quite consistent. As labor and management interact in a given setting over time, each develops a somewhat stable frame of reference regarding the orientations of the other party; in effect, this frame of reference describes where the fundamental relationship lies along a cooperation-conflict continuum. Within such frames of reference the parties engage in sensemaking behaviors—interpreting the meanings of otherwise ambiguous events and making attributions (Shaver, 1975) regarding the intentions, motivations, and beliefs that underlie the other party's actions.

It appears reasonable to expect that, whatever other variables might influence the likelihood of dual commitment, the ambient relationship climate will be a factor. Social-psychological theories of cognitive consistency predict, for example, that individuals who try to be loyal to two conflicting social systems will encounter considerable cognitive dissonance (Festinger, 1957). Indeed, both the old and the new research on dual allegiance or dual commitment contain this premise. Gottlieb and Kerr (1950), in their survey of the Buchsbaum Company, concluded that the high correlation between attitudes toward management and toward the union was associated with a 62-year history that was devoid of union-management hostility. Rosen (1954) also suggested that dual allegiance is most likely under circumstances in which no intense role conflict is activated. Furthermore, although Dean (1954) concluded that "there is little . . . that precludes dual loyalty, even where the union-management relationship is antagonistic" (1954: 536), the patterns she reported in her comparison of three plants appeared to indicate that participation in union activities and dual allegiance were more apt to occur together in the plant whose climate she designated cooperative than in the plants whose relationships were characterized by conflict. Her observation suggests that the intensity of members' personal involvement in a union might moderate any relationship between dual loyalty and labor-management relationship climate.

In more recent research, Martin (1981), using Walker and Lawler's (1979) classification scheme, found that dual allegiance was typical of union members in protective unions, and unilateral allegiance was typical in aggressive unions. Fukami and Larson (1982) found a significant correlation between commitment to union and to company ($r = .27$, $p < .001$). However, when they partialled out employees' perceptions of the union-management relationship, the correlation fell to a nonsignificant .06. Gallagher (1984) reported a similar finding; when he controlled for individual perceptions of the quality of the union-management relationship, positive and statistically significant relationships between union members' organizational commitment and union commitment became nonsignificant.

By and large, the research we have reviewed—particularly the new wave of research—tends to lead to the following conclusions: (1) dual commitment is a relatively common phenomenon—70 percent of Gallagher's (1984) union members were so classified; (2) the ambient labor-management relationship climate might explain much of the variance in dual commitment. However, both the old and the new studies have been vulnerable to certain response artifacts. It has been the norm to assess dual commitment by counting the numbers of people whose attitudes were positive toward both union and management. Furthermore, in each of the studies reported earlier in this paper, both the commitment measures and the judgments of the quality of the overall union-management relationship derived from the same self-reports. Thus, yea-saying response bias could inflate both the extent of dual commitment per se and observed covariation in measures of relationship climate and commitment.

Overcoming such artifacts necessitates deriving a direct measure of dual commitment. It also requires using independent sources to measure: (1) labor-management relationship climate, and (2) employee commitments. Providing an opportunity for variance to occur in a relationship-climate measure mandates conducting a study across several independent units or organizations rather than in a single organization that would theoretically have a single climate.

The research reported here investigated the relationships between ambient labor-management relationship climates and both dual and separate member commitments to employing organizations and unions. In addition, we assessed the moderating influence of member participation in union activities on this relationship. We included this aspect in view of Dean's (1954) observation that dual loyalty and active participation in union activities tended to co-occur where an overall relationship was relatively cooperative.

Hypotheses were as follows:

> *Hypothesis 1: Relatively high levels of dual commitment will occur in organizations whose ambient labor-management relationship climates are relatively cooperative.*
>
> *Hypothesis 2: The relative level of members participation in their unions' activities will moderate the relationship between climate and dual commitment. This relationship will be stronger under conditions of high member participation than under low participation.*

In addition to investigating the hypothesized relationships, an exploratory analysis investigated the covariance between the labor-management relationship climate and sole commitments to organization or union. However, we established no directional hypotheses concerning sole commitments. Although observers often note that conflict with an out-group polarizes in-group members and increases their cohesiveness, in this instance the members in question were not members of one system alone but were attached to both groups. Thus our treatment remained exploratory.

METHODS

Setting and Participants

This research was part of a larger study of labor-management relations and organizational performance in 28 separate municipal bus companies in the western United States. We gained access to each organization by corresponding with its general manager. We obtained access to union members through separate negotiations with the leadership of the appropriate unions.

Teams of two researchers made two-day long visits to sites in the summer of 1977. We administered separate-form questionnaires to: (1) transit organization managers who were involved in union-management relations; (2) their opposite-number labor leaders; and (3) rank-and-file employees who were members of the bus drivers' bargaining unit. As we were able to administer the rank-and-file questionnaires in only 22 of the 28 organizations, the organizational *N* for purposes of this article is 22.

With the number of bus drivers employed serving as an indicator, organizations varied in size from 13 to 4,940, with a median size of 58. When maintenance personnel, who were normally in the drivers' bargaining unit and thus included in the survey, are taken into account, median organization size was 69.

Eighteen of the organizations were unionized, 11 in transit unions and seven in general unions; drivers in the remaining four were represented in collective bargaining by local employee associations that were functional equivalents of unions. Although only 12 of the organizations had union shops and three others were agency shops, 85 percent of the respondents to the rank-and-file survey were members of unions or associations. During their most recent rounds of negotiations, strike votes had occurred in eight of the organizations and the membership had supported strikes in seven; however, only one organization had seen an actual strike.

Respondents either hand-delivered their anonymous completed questionnaires to us or mailed them in prepaid mailer envelopes. Response rates and usable response Ns were 84 percent ($N=67$) for managers, 73 percent ($N=57$) for labor leaders, and 64 percent ($N=1{,}244$) for rank-and-file employees. Only members of the union or employee association in each organization answered questions on dual and union commitment. This reduced the rank-and-file N for analyses involving union commitment or dual commitment to 1,057.

Bus drivers accounted for 89 percent of the rank-and-file sample; 6 percent were driver supervisors or bus maintenance personnel; the remaining 5 percent were clerical and related personnel. Rank-and-file respondents were 82 percent men, 74 percent white, 14 percent black, and 10 percent Hispanic. Mean age was about 38, with 28 percent younger than 30 and 18 percent older than 50. Eighteen percent had worked less than two years in their present organizations, and 30 percent had been employed there longer than five years.

Measures

Each question in the scales listed in this section had a 7-point Likert agree/disagree format. We incorporated scales as appropriate in questionnaires administered to transit managers, labor leaders, and rank-and-file employees. The Appendix contains all items for these scales, except for the Organizational Commitment Questionnaire, which has been fully described elsewhere (Mowday et al., 1979).

Manager and Labor Leader Questionnaires

Labor-management relationship climate. The ambient labor-management relationship climate for each organization was measured by a 23-item scale. We adapted most of the items from the University of Michigan Organizational Assessment Package (Institute for Social Research, 1975). Parallel forms were developed so that each side responded to questions about their perceptions of their opposite numbers. Reliabilities for the scales (α) were .92 for managers and .95 for labor leaders.

Employee Questionnaire

Dual commitment. A 5-item Likert scale measured dual commitments to employing organizations and to unions or employee associations ($\alpha = .71$).

Commitment to organization. The 15-item Organizational Commitment Questionnaire (OCQ) (Mowday et al., 1979) measured rank-and-file employees' commitments to their employing organizations ($\alpha = .89$).

Commitment to union. A 4-item scale measured members' commitments to unions or employee associations ($\alpha = .76$).

Union participation. We asked union or association members eight yes or no questions regarding participation in various union activities in order to derive their relative levels of union participation. Each employee's union participation score was a simple sum of yes answers.

Analysis

Including several separate organizations in the research and employing separate questionnaires for managers, labor leaders, and rank-and-file union members enabled use of independent data sources for measuring independent and dependent variables. We assessed labor-management relationship climates by means of the managers' and union leaders' questionnaires. Dual commitment, sole commitments to union or employing organization, and union participation were measured through questionnaires administered to the rank-and-file. This approach eliminated the problem of single-source contamination of independent and dependent variables for all analyses except those relating union participation to commitment.

Administration of transit-manager questionnaires was limited to general managers and other managers who had been directly involved in labor-management relations. We gave the labor-leader questionnaires to their counterparts on the union side. Accordingly, the measure of each organization's ambient labor-management relationship climate derived from reports of persons best able to make that assessment. This procedure resulted in our obtaining from 4 to 11 usable returns per organization.

To derive scores for labor-management relationship climate for each organization, we first aggregated the scores within each organization for (1) labor-leaders and (2) managers. We then assigned the mean of these two scores as an organization's score on labor-management relationship climate. All rank-and-file members then received one of 22 climate scores that were based on the consensus of the labor leaders and managers from their own organizations.

Constructing organization-level variables by aggregating individual perceptions is problematic, but under the right conditions this approach can yield valid indices of organizational properties (Lincoln & Zeitz, 1980). In order to justify our assigning a single climate score to each organization, we considered it essential that there be high degrees of interrater agreement within each organization. In particular, union and management sources should report highly similar perceptions of the labor-management climate at given locations. Historically, however, the evidence for such consensus has been mixed. Rim and Mannheim (1964) found a significant correlation ($r = .30$, $p < .05$) between selected labor-management attitudes of general managers and union representatives. But Stagner, Chalmers, and Derber (1958) concluded from the Illini City data that relationships above the level of the rank-and-file are too complex to permit consensus regarding a labor-management relationship climate.

In the present research, the Pearson product-moment correlation of these pairs of scores for the 22 organizations was .84 ($p < .001$). To ascertain whether this ostensible agreement might have resulted from a washing out of within-group variance through the averaging process, we conducted a one-way analysis of variance across the 22 organizations, with the identities of the organizations involved treated as the categorical variable. The individual scores of all labor and management respondents comprised the dependent variable. The F-ratio obtained was 6.25 ($p < .0001$). We repeated this analysis twice: once using only labor leaders' scores, and once using only managers' scores. Again, in each instance, the F-ratio was statistically significant ($p < .001$).

Hierarchical regression analysis was performed in order to assess the relationship between the labor-management climate and dual commitment and to test for the moderating influence of union participation. We followed Arnold's (1982) procedural recommendations and used dual commitment as the dependent variable. This procedure was designed not only to measure each independent variable's unique influences on dual commitment, but also to test for the presence of a true moderator variable that altered the form of the relationship between dependent and independent variables. As we had hypothesized that the main effect would be the relationship between climate and dual commitment, and that union participation would moderate this effect, we entered three independent variables into the regression in the following order: (1) labor-management climate, (2) union participation, and (3) a multiplicative interaction term (climate × participation).

RESULTS

Table 1 presents descriptive statistics and the intercorrelation matrix for the variables of interest in this research. For the overall sample, the correlation between labor-management relationship climate and dual commitment was .32 ($p < .001$). Although, not surprisingly, the correlation between union participation and commitment to a union was statistically significant ($r = .22$, $p < .001$), the correlations between union participation and dual commitment and between union participation and organizational commitment were approximately zero.

There was no apparent relationship between labor-management relationship climate and recent strike-related activity. The one organization that had experienced a strike was near the median of the labor-management climate rank ordering, and the seven that had voted to strike were distributed nearly evenly across three-quarters of the range of scores.

Figure 1 portrays the overall relationship between labor-management climate and dual commitment. Each data point represents one organization, plotted at its members' mean level of dual commitment on the ordinate and at the organization's score for labor-management relationship climate along

TABLE 1
Descriptive Statistics and Intercorrelation Matrix for Commitment, Climate, and Participation[a]

Variables	Means	s.d.	1	2	3	4
1. Dual commitment	4.39	1.21				
2. Union commitment	4.65	1.27	.40*** (1,035)			
3. Organizational commitment	4.50	1.12	.48*** (1,051)	.30*** (1,052)		
4. Labor-management relationship climate	4.53	0.73	.32*** (1,057)	.09** (1,056)	.19*** (1,205)[b]	
5. Union participation	2.48	1.20	.03 (1,038)	.22*** (1,027)	−.02 (1,076)	.07* (1,098)

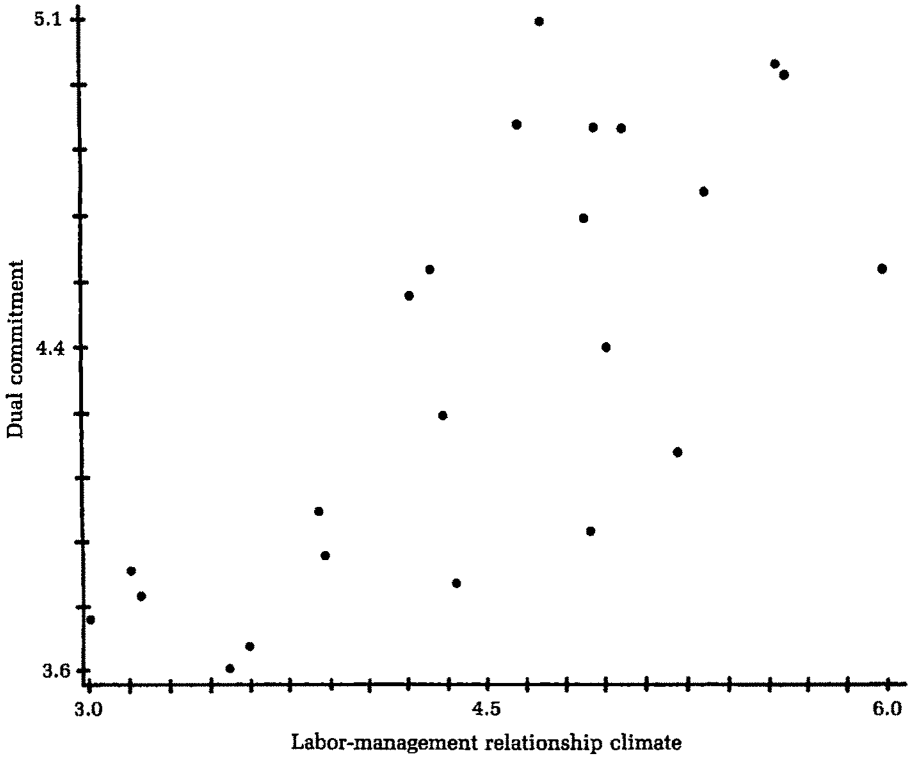
[a] Figures below each correlation coefficient in parentheses indicate N for the correlation.

[b] Organizational commitment measure included employees who were not union members.

[c] Mean and s.d. are based on 1205 cases, each of which was assigned one of 22 possible labor-management climate scores. These scores were derived from 155 responses of managers and labor leaders; the mean of that distribution was 4.48, with a standard deviation of 1.10.

* $p<.05$
** $p<.01$
*** $p<.001$

FIGURE 1
Scatterplot of the Labor-Management Relationship Climate and Dual Commitment[a]



[a] Each plotted point is the average pair of scores for one organization.

the abscissa. The correlation between dual commitment and climate, with organizations as the units of analysis, is .73 ($p <.001$).

Table 2 summarizes the results of the moderated regression analysis. The labor-management relationship climate explains about 10.3 percent of the variance in dual commitment ($p < .001$). When union participation is entered next, there is virtually no change in R^2 and the F-to-enter is a nonsignificant 0.1. Finally, when the interaction term, climate × participation, is entered, explained variance climbs to 13.4 percent ($p < .001$). Thus, union participation appears to meet the test of a true moderator variable; it accounts for a significant increment in variance explained in interaction with the designated independent variable, labor-management relationship climate (Arnold, 1982).

TABLE 2
Hierarchical Multiple Regression Testing
the Moderating Effect of Union Participation on Dual Commitment

Steps/Variables Entered	B	F^a	p	R^2 Change	Cumulative R^2
1. Labor-management relationship climate	.535	118.4	.000	.103	.103
2. Participation	.011	0.1	n.s.	.000	.103
3. Interaction term (1 × 2)	.054	36.4	.000	.031	.134

[a] F to enter or remove.

The effect of the interaction term, which had a positive regression coefficient, was to amplify the relationship between the labor-management relationship climate and dual commitment; the relationship covaried with activity. To illustrate this effect, we performed extreme-groups regressions for subsamples one standard deviation above and one standard deviation below the mean for union participation. Whereas the regression coefficient for climate on dual commitment had been .535 for the entire sample, it increased to .603 for the high-participation subgroup and decreased to .446 for the low-participation subgroup.

Because there were large variations in organizations' sizes, we were concerned that either the labor-management relationship climate or its relationship with dual commitment might somehow be a function of size. However, the Pearson correlation between size and labor-management climate was only -.03 (n.s.). Furthermore, the Pearson correlation between dual commitment and the labor-management relationship climate ($r = .32$) was not changed by partialling out organization size.

In order to illustrate the monotonic nature of the relationship between the labor-management relationship climate and dual commitment, a subsidiary analysis was conducted. We divided the sample of organizations into subgroups and performed a one-way analysis of variance with dual commitment as the dependent variable. For this purpose we arbitrarily selected five climate categories, though the point could be illustrated with fewer or more categories.[2] We partitioned the rank-and-file sample into categories by dividing it at the 20th, 40th, 60th and 80th percentile points of the distribution for labor-management relationship climate.

Table 3 illustrates the results of that analysis. Dual commitment was strongest in organizations and bargaining units that coexisted in the most cooperative relationship climate and weakest in the least cooperative climate, with a nearly linear progression in between. However, only two of the four intercell contrasts were statistically significant.

---

[2] Walton and McKersie's (1965) well-known scheme contains five categories, although we make no claim that our subgroups accurately reflect their classifications.

TABLE 3
ANOVA Results: Dual Commitment and Sole Commitments to Union and Management with Labor-Management Relationship Climate as the Categorical Variable

Independent Variables	Category (N=)[a,b]	Means	F	Cell Contrasts
Dual commitment	1 (142)	4.83	30.72***	
	2 (333)	4.66		1–2
	3 (321)	4.41		2–3**
	4 (32)	4.23		3–4
	5 (229)	3.70		4–5*
Organizational commitment	1 (159)	4.89	11.74***	
	2 (374)	4.62		1–2*
	3 (347)	4.31		2–3***
	4 (34)	4.63		3–4
	5 (266)	4.26		4–5
Union commitment	1 (145)	4.49	8.60***	
	2 (335)	4.96		1–2***
	3 (317)	4.60		2–3***
	4 (28)	4.19		3–4
	5 (230)	4.44		4–5

[a] Labor-management relationship climate, the categorical variable, was divided into 5 categories at the 20th, 40th, 60th, and 80th percentiles. The most cooperative climate is category 1 and the least cooperative is category 5.

[b] Union commitment and dual commitment Ns include only union members; organizational commitment N includes all rank-and-file employees.

*p<.05

**p<.01

***p<.001

Table 3 also contains equivalent one-way analyses of variance of organizational commitment and union commitment across the same five categories of labor-management relationship patterns. Although commitment to employing organizations followed a pattern similar to that of dual commitment, there was a slight departure from monotonicity in one cell. In the case of union commitment, a sharp departure from monotonicity occurred, with the extreme groups showing nearly identical levels of commitment to unions. Whereas dual commitment had shared 10.2 percent common variance with the labor-management relationship pattern, common variance was only 3.6 percent for organizational commitment and less than 1 percent for union commitment. We performed no moderated regression analyses for these two independent variables because the analyses were exploratory; interpretation of any significant interaction effect is legitimate only where there has been an a priori theory suggesting such an effect (Arnold, 1982).

DISCUSSION

Using independent measures of labor-management relationship climate and dual commitment, this study confirmed that the extent of dual commitment expressed by rank-and-file members is related to the ambient relationship between labor and management within their organizations. The likelihood of simultaneous commitment to two interacting systems such as a union and an employing organization appears to grow where the relationship between the systems is cooperative. Although this finding is both intuitively reasonable and congruent with some of the earlier research cited in this paper, we are unaware of any other research on this topic in which data have so thoroughly supported theory.

The finding of a significant interaction between the quality of a union-management relationship and the extent of members' union participation is consistent with our fundamental argument. We stated that simultaneous commitment to two social systems is problematic to the extent that those systems are not cooperative and may force members to make either-or choices. To review the basic premise, we posited that an ambient relationship climate would establish the extent to which members would face choice dilemmas in forming loyalty attachments to the two systems. Yet it has long been known that people are imperfectly rational and can often compartmentalize their logic so as to hold simultaneously incompatible opinions (Cyert & March, 1963; Simon, 1976). That is, they may be able to do so until they must confront inherent incompatibilities. The effect of member participation, then, is to make any existing incompatibility salient and thus to amplify the fundamental effect of climate on dual commitment. In effect, where two systems tend toward the low-cooperation end of the labor-management climate continuum, the inherent choice dilemma is most salient to union activists.

In combination, these two variables—labor-management relationship climate and union participation—seem to go a long way toward clarifying the mixed findings of earlier research on dual allegiance that failed to resolve the question of whether workers can be loyal to both their employing organizations and their unions. Perhaps earlier researchers were asking the wrong question. A more appropriate question might be, "What are the factors that enhance or decrease the likelihood that workers will develop dual commitment?" The research reported here has provided at least a partial response to that question by implicating labor-management relationship climate and member participation in union affairs as factors that influence the extent of dual commitments.

Some methodological observations are in order. We admit that measuring an organizational attribute by aggregating self-reported data from members, as we did to create the variable for labor-management relationship climate, is controversial. Indeed, the study of organizational climate has endured much debate as to whether there is any such thing as an organizational climate (James & Jones, 1974; Schneider, 1975). The criticism is essentially that researchers using survey methods have simply combined disparate perceptions into fallacious averages.

In this instance, however, we are confident that our measure of the labor-management relationship climate represents a meaningful attribute of each organization, not a spurious consensus. The scale's reliabilities were themselves high. Moreover, the high correlation between the perceptions of managers and their labor counterparts indicated that labor and management appeared to see ambient labor-management relationship climates in much the same way. One-way analyses of variance showed that within-organization variance on perceptions of climate was relatively low when compared to between-organization variance. Moreover, the uniformly strong relationships between these managers and labor leaders' perceptions of climate and the commitment levels of other members of the same organizations add further evidence that labor-management relationship climates, as measured in this study, have real meaning. In sum, we believe that the data justify our treating labor-management relationship climate as an organizational attribute.

Regarding causality, cross-sectional research of this type obviously cannot establish whether ambient labor-management relationship climates actually result in varying levels of dual commitment or whether other factors codetermine climates and commitments. We believe a logical case can be made that role dilemmas imposed by membership in conflicting systems must necessitate a loyalty choice between adversaries. However, it remains to more elaborate, perhaps longitudinal, designs to have the final say.

Assuming the correctness of the causal relationship identified in the preceding discussion, we see these findings as having important implications for both labor and management, especially in light of current contextual developments. Deregulation of heavily unionized sectors—airlines, trucking—and increasing internationalization of competitive markets are two examples of such developments. Adapting to these changes will increasingly require both parties to make strategic choices about appropriate mixes of conflict and cooperation in their relations. This research suggests that both substantive strategies concerning extent of conflict or cooperation and unions' processes for implementing them will bear on members' subsequent attachments to unions and to employing organizations. These attachments are, in turn, of consequence because of their relationships to organizational performance. For example, previous research has shown that decertification is more likely for unions where dual commitment is low (Angle & Perry, 1984). Leadership efforts to improve a climate, rather than to hold firmly to an adversarial position, would appear to be the only rational strategic choice, particularly when the survival of one or both parties is problematic and members' loyalties are essential. Although this implication from the research is somewhat speculative, it is ripe for additional study.

Dual commitment to unions and employing organizations has been studied only in the context of ongoing labor-management relationships. It may be equally or more important in relation to union efforts to organize new employers. The labor movement's current organizing difficulties give some credibility to this view. Employees' perceptions that labor is too adversarial are among the causes to which these difficulties have been traced. Employees' tolerances for role dilemmas imposed by membership in conflicting systems and their expectations about the likelihood of encountering such dilemmas influence their votes in union representation elections. A no-union vote may reflect employees' preferring to avoid the problem of divided loyalties. Again, if this conjecture is correct, it implies that it will be necessary for unions to change their adversarial images and to develop cooperation-oriented strategies.

Bakke (1946) delivered the message that unions and management share a superordinate goal of mutual survival. As Whyte (1949) pointed out, labor unions and their management counterparts comprise "two social systems made up of highly interdependent parts . . . one is mutually dependent upon the other" (1949:13). In a similar vein, Drucker (1949) predicted that management and unions might someday destroy each other unless they could establish conditions under which their shared members could be loyal to one without having to abandon loyalty to the other. We believe that the data reported here substantiate the views of these eminent scholars.

REFERENCES

Angle, H. L., & Perry, J. L. 1983. Organizational commitment: Individual and organizational influences. ***Work and Occupations***, 10: 123–146.

Angle, H. L., & Perry, J. L. 1984. Union member attitudes and bargaining unit stability in urban transit. In B. Dennis (Ed.), ***Proceedings of the 36th Annual Meeting of the Industrial Relations Research Association***. Madison, Wis.: IRRA.

Arnold, H. 1982. Moderator variables: A clarification of conceptual, analytic, and psychometric issues. ***Organizational Behavior and Human Performance***, 29: 143–174.

Bakke, E. W. 1946. ***Mutual survival: The goal of union and management***. New York: Harper.

Barkin, S. 1950. A trade unionist appraises management personnel philosophy. ***Harvard Business Review***, 28 (5): 59–64.

Becker, H. S. 1960. Notes on the concept of commitment. ***American Journal of Sociology***, 66: 32–40.

Buchanan, B., II. 1974. Building organizational commitment: The socialization of managers in work organizations. ***Administrative Science Quarterly***, 34: 339–347.

Chalmers, W. E., Chandler, M. K., McQuitty, L. L., Stagner, R., Wray, D. E., & Derber, M. 1954. ***Labor-management relations in Illini City***, vol. 2. Champaign, Ill.: Institute of Labor and Industrial Relations, University of Illinois.

Cyert, M., & March, J. G. 1963. ***A behavioral theory of the firm***. Englewood Cliffs, N. J.: Prentice-Hall.

Dean, L. R. 1954. Union activity and dual loyalty. ***Industrial and Labor Relations Review***, 7: 526–536.

Drucker, P. F. 1949. ***The new society: The anatomy of industrial order***. New York: Harper & Row Publishers.

England, G. W. 1960. Dual allegiance to company and union. ***Personnel Administration***, 23(2): 20–25.

Etzioni, A. 1975. ***A comparative analysis of complex organizations*** (rev. & enl. ed.). New York: Free Press.

Ferris, K. R., & Aranya, N. 1983. A comparison of two organizational commitment scales. ***Personnel Psychology***, 36: 87–98.

Festinger, L. A. 1957. ***A theory of cognitive dissonance***. Stanford, Calif.: Stanford University Press.

Fukami, C. V., & Larson, E. 1982. ***The relationship between union commitment and organizational commitment: Dual loyalty reexamined.*** Paper presented at the 42nd Annual Meeting of the Academy of Management, New York.

Fukami, C. V., & Larson, E. 1984. Commitment to company and union: Parallel models. ***Journal of Applied Psychology***, 69: 367–371.

Gallagher, D. G. 1984. The relationship between organizational and union commitment among federal government employees. ***Proceedings of the Academy of Management***: 319–323.

Gordon, M. E., Philpot, J. W., Burt, R. E., Thompson, C. A., & Spiller, W. E. 1980. Commitment to the union: Development of a measure and an examination of its correlates. ***Journal of Applied Psychology***, 65: 479–499.

Gottlieb, B., & Kerr, W. A. 1950. An experiment in industrial harmony. ***Personnel Psychology***, 3: 445–453.

Gouldner, A. W. 1957. Cosmopolitans and locals: Toward an analysis of latent social roles. ***Administrative Science Quarterly***, 2: 281–306.

Harbison, F., & Coleman, J. R. 1951. ***Goals and strategy in collective bargaining***. New York: Harper.

Institute for Social Research. 1975. ***Michigan organizational assessment package: Progress report II***. Ann Arbor, Mich.: Survey Research Center, ISR.

James, L. R., & Jones, A. P. 1974. Organizational climate: A review of theory and research. ***Psychological Bulletin***, 81: 1096–1112.

Kahn, R. L., Wolfe, D. M., Quinn, R. P., Snoek, J. D., & Rosenthal, R. A. 1964. ***Organizational stress: Studies in role conflict and ambiguity***. New York: John Wiley & Sons.

Kanter, R. M. 1972. ***Commitment and community: Communes and utopias in perspective***. Cambridge, Mass.: Harvard University Press.

Kelman, H. C. 1958. Compliance, identification, and internalization: Three processes of attitude change. ***Journal of Conflict Resolution***, 2: 51–60.

Kiesler, C. A. 1971. ***The psychology of commitment: Experiments linking behavior to belief***. New York: Academic Press.

Kornhauser, A., Sheppard, H. L., & Mayer, A. J. 1956. ***When labor votes***. New York: University Books.

Ladd, R. T., Gordon, M. E., Beauvais, L. L., & Morgan, R. L. 1982. Union commitment: Replication and extension. ***Journal of Applied Psychology***, 67: 640–644.

Lieberman, S. 1956. The effects of changes in roles on the attitudes of role occupants. ***Human Relations***, 9: 385–402.

Lincoln, J. R., & Zeitz, G. 1980. Organizational properties from aggregate data: Separating individual and structural effects. ***American Sociological Review***, 45: 391–408.

Martin, J. E. 1981. Dual allegiance in public sector unionism: A case study. ***International Review of Applied Psychology***, 30: 245–259.

Martin, J. E., Mageneau, J. M., & Peterson, M. F. 1982. Variables differentiating patterns of commitment among union stewards. ***Proceedings of the Academy of Management***: 302–306.

Miller, G. A., & Wager, L. W. 1971. Adult socialization, organizational structure, and role orientations. ***Administrative Science Quarterly***, 16: 151–163.

Mowday, R. T., Steers, R. M., & Porter, L. W. 1979. The measurement of organizational commitment. ***Journal of Vocational Behavior***, 14: 224–247.

Mowday, R. T., Steers, R. M., & Porter, L. W. 1982. ***Employee-organization linkages: The psychology of commitment, absenteeism, and turnover***. New York: Academic Press.

Porter, L. W., Steers, R. M., Mowday, R. T., & Boulian, P. V. 1974. Organizational commitment, job satisfaction, and turnover among psychiatric technicians. ***Journal of Applied Psychology***, 59: 603–609.

Purcell, T. V. 1960. ***Blue collar man: Patterns of dual allegiance in industry***. Cambridge, Mass.: Harvard University Press.

Rim, Y., & Mannheim, B. F. 1964. Factors related to attitudes of management and union representatives. ***Personnel Psychology***, 17: 149–165.

Ritzer, G., & Trice, H. M. 1969. ***An occupation in conflict: A study of the personnel manager***. Ithaca, N.Y.: New York State School of Industrial and Labor Relations, Cornell University.

Rizzo, J. R., House, R. J., & Lirtzman, S. I. 1970. Role conflict and ambiguity in complex organizations. ***Administrative Science Quarterly***, 15: 150–163.

Rosen, H. 1954. Dual allegiance: A critique and a proposed approach. ***Personnel Psychology***, 7: 67–71.

Salancik, G. R. 1977. Commitment and the control of organizational behavior and belief. In B. M. Staw & G. R. Salancik (Eds.) ***New directions in organizational behavior:*** 1–54. Chicago: St. Clair.

Schneider, B. 1975. Organizational climates: An essay. ***Personnel Psychology***, 28: 447–479.

Shaver, K. 1975. ***An introduction to attribution processes***. Cambridge, Mass.: Winthrop.

Simon, H. A. 1976. ***Administrative behavior*** (3rd ed.). New York: Macmillan Book Publishing Co.

Stagner, R., Chalmers, W. E. & Derber, M. 1958. Guttman-type scales for union and management attitudes toward each other. ***Journal of Applied Psychology***, 42: 293–300.

Staw, B. M. 1977. ***Two sides of commitment***. Paper presented at the 37th Annual Meeting of the Academy of Management, Orlando, Fla.

Stevens, J. M., Beyer, J. M., & Trice, H. M. 1978. Assessing personal, role, and organizational predictors of managerial commitment. ***Academy of Management Journal***, 21: 380–396.

Walkers, J. M., & Lawler, J. J. 1979. Dual unions and political processes in organizations. ***Industrial Relations***, 18: 32–43.

Walton, R. E., & McKersie, R. B. 1965. ***A behavioral theory of labor negotiations: An analysis of a social interaction system***. New York: McGraw-Hill Book Co.

Whyte, W. F. 1949. Patterns of interaction in union-management relations. ***Human Organization***, 8 (4): 13–19.

APPENDIX
Scales and Included Items

Labor-management relationship climate, managers' form[a]

In general:

1. The union is reasonable in dealing with management.
2. The union weakens employee discipline.[b]
3. The union will give in to management when management is right.
4. The union deals openly with management.
5. The union abuses its power.[b]
6. The union tries to cooperate with management.
7. The union doesn't understand management's problems.[b]
8. Management's relations with the union are satisfactory.
9. The union and management are natural enemies.[b]
10. In this organization, relations with the union are better than they used to be.
11. I expect a strike at this organization within the coming year.[b]
12. There are many *rules* which tend to *dictate* the settlement reached on specific issues[c] .

In negotiations:

13. The union uses pressure tactics.[b]
14. The union and management are hostile toward each other.[b]
15. The union makes concessions to avoid problems.
16. The union won't listen to new ideas.[b]
17. The union and management share most information.
18. The union won't give in on anything unless they are forced.[b]
19. Both sides are willing to try new solutions.
20. The union and management work together to try to find creative solutions to problems.

In dealing with grievances:

21. The union tries to understand management's side.
22. The union uses pressure tactics.[b]
23. The union is more interested in supporting its members than in "what is right."[b]
24. Management shares relevant information with the union.[c]
25. The union shares all relevant information with management.

Labor-management relationship climate, union form

The union form was an adaptation of the management form. We reworded 24 of the items in a straightforward manner so that union leaders could indicate their perceptions of management. Example: Management is reasonable in dealing with the union. For item 2, however, simple rewording was not feasible. The item was rewritten as: Management tries to interfere in internal union affairs.[b]

Union commitment

1. I am loyal to the union.
2. I talk up this union to my friends as a great union to belong to.
3. I find it hard to agree with my union's policies.[b]
4. I don't care if the union survives.[b]

Dual commitment

1. It is easy to be loyal to both the union and management.
2. The management makes it easy to conduct union business.
3. The management makes it difficult for me to talk to my group representative or job steward.[b]
4. You can't be a union member and support management at the same time.[b]
5. The union helps me deal effectively with management.
6. Union members don't like it if you try to help management improve work effectiveness.[c]

Union participation

Are you currently serving, or during the last five years, have you served in any of the following union offices?

1. Local president, secretary, treasurer, etc.
2. Local steward or group representative.
3. Local member of a committee.
4. District elected official.
5. District appointed committee.
6. Did you vote in the last union election?
7. Have you read the union contract?
8. Do you read the union newsletter?
9. Have you filed a grievance in the past two years?

---

[a] Adapted from the University of Michigan Organizational Assessment Package.

[b] Reverse-scored item.

[c] Item deleted from scale to maximize reliability.

Harold L. Angle is Associate Professor of Strategic Management and Organization in the School of Management, University of Minnesota. He received his Ph.D. degree from the University of California, Irvine. His current research interests include attachment to work and work organizations, psychological contracts, and organizational innovation.

James L. Perry is Professor of Public and Environmental Affairs at Indiana University. He received his Ph.D. degree from the Maxwell School, Syracuse University. His current research interests include organizational structure and performance in the public sector.

© *Academy of Management Journal*
1986, Vol. 29, No. 1, 51–71.

GROUP APPROACHES FOR IMPROVING STRATEGIC DECISION MAKING: A COMPARATIVE ANALYSIS OF DIALECTICAL INQUIRY, DEVIL'S ADVOCACY, AND CONSENSUS

DAVID M. SCHWEIGER
WILLIAM R. SANDBERG
JAMES W. RAGAN
University of Houston

This laboratory study compared the effectiveness of the dialectical inquiry, devil's advocacy, and consensus approaches to strategic decision making by groups. Results showed that both dialectical inquiry and devil's advocacy led to higher quality recommendations and assumptions than consensus. Dialectical inquiry was also more effective than devil's advocacy with respect to the quality of assumptions brought to the surface. However, subjects in the consensus groups expressed more satisfaction and desire to continue to work with their groups and greater acceptance of their groups' decisions than did subjects in either of the two other types of group studied.

Top managers are often required to anticipate, define, and solve "wicked" strategic problems. Several elements characterize these complex problems: (1) numerous complicated linkages among organizational and environmental elements; (2) dynamic and uncertain environments; (3) ambiguity of available information; (4) lack of complete information; and (5) conflicts concerning the outcomes of decisions among interested parties (Mason & Mitroff, 1981). These characteristics produce decision environments in which single right recommendations are rarely apparent. Such complexity often forces top managers to draw inferences and assumptions about their organizations and environments from available information and then to try to define and solve problems.

Usually groups of managers with a variety of information and different perspectives address strategic problems (Brodwin & Bourgeois, 1984). Discussion and other interaction among top executives are frequent means of sharing and evaluating information and ensuing inferences, assumptions, and recommendations (Glueck, 1980; Mintzberg, Raisinghani, & Theoret, 1976; Springer & Hofer, 1978; Stagner, 1969). The group process may also be important in gaining commitment to and acceptance of decisions among individuals who ultimately will be responsible for implementing such decisions (Mason & Mitroff, 1981).

Group processes themselves do not ensure that top managers will adequately explore available information. Groups frequently avoid uncertainty and prematurely smooth over conflict (Brodwin & Bourgeois, 1984). Janis (1972) clearly described the problems inherent in consensus-seeking groups that value group harmony more highly than critical evaluation. The nature of interactions and the processes by which information is shared and evaluated appear to be critical factors in effective strategic decisions by top management groups. Thorough evaluations and critiques of inferences and assumptions are essential (Janis, 1972). The inferences and assumptions that survive such an evaluation and critique are more likely to be valid than those that do not survive (Mitroff, 1982b; Mitroff & Mason, 1981). Without such a rigorous process it is probable that decision makers will retain intuitively obvious but invalid assumptions that will reduce the quality of groups' solutions and final recommendations.

During the past few years, there has been much debate and comparative research on the best approach for structuring strategic decision making groups. Researchers have paid the most attention to two approaches that rely on the use of constructive group conflict, dialectical inquiry and devil's advocacy (Cosier, 1978, 1980; Cosier & Aplin, 1980; Cosier, Ruble, & Aplin,1978; Lourenco & Glidewell, 1974; Mitroff, Barabba, & Kilmann, 1977; Schwenk, 1982b, 1984; Schwenk & Cosier, 1980). Dialectical inquiry uses debates between diametric sets of recommendations and assumptions, whereas devil's advocacy relies on critiques of single sets of recommendations and assumptions. In spite of the amount of research on these approaches, debate continues concerning the validity of the findings from these studies and the conclusions that can be drawn from them (Cosier, 1981, 1982, 1983; Mitroff, 1982a, 1982b; Mitroff & Mason, 1981; Schwenk, 1982a, 1982c, 1983).

On the basis of two field studies (Lourenco and Glidewell, 1974; Mitroff et al., 1977), Mitroff and his colleagues contended that dialectical inquiry is the best approach available for improving strategic decisions. In turn, Cosier and his associates argued that these two studies have not been able to establish a causal relationship between this approach and decision performance. On the basis of a series of laboratory experiments, they concluded that devil's advocacy may be more effective than dialectical inquiry. Although these laboratory studies have permitted causal interpretation, a recent study by Schweiger and Finger (1984) suggested that inadequate experimental tasks, conceptualizations, and operational definitions of the two approaches may have prevented these studies from capturing their essences. These studies relied on such tasks as the Multi-Cue Probability Learning Paradigm, a creative problem-solving task, and case studies. All of these studies required otherwise naive individuals to make decisions or judgments after either written or videotaped exposure to contrived recommendations and supporting assumptions made by perceived experts. In the case of dialectical inquiry, studies presented plans and counterplans; for devil's advocacy, plans and critiques. These research efforts studied neither approach as a group process nor were there any opportunities for subjects to participate in development of recommendations, assumptions, or critiques. The conceptualizations and operational definitions used in previous laboratory studies were at variance with those originally proposed by Mitroff, Mason, and their colleagues (Mason, 1969; Mason & Mitroff, 1981; Mitroff, 1982a, 1982b; Mitroff & Emshoff, 1979).

An examination of how the two approaches were intended to function will illustrate weaknesses in previous laboratory research. Both dialectical inquiry and devil's advocacy recommend the use of constructive conflict in group decision making. Formalized argumentation and debate among top managers allow assumptions and recommendations to be systematically evaluated and the strengths and weaknesses of each to be made explicit (Mason & Mitroff, 1981). Therefore, an adequate assessment of the effectiveness of the two approaches requires a research design that incorporates active interaction among participants in decision making groups. As Mitroff succinctly noted:

> the dialectical inquiry method [and devil's advocate method are] . . . above all . . . *learning* processes whereby through active, heated, and intense debate between and among *interested* parties (not naive subjects) the parties came to discover and to invent entirely new alternatives as well as elaborate on old ones (1982b: 222).

Although both approaches employ conflict as their primary structural mechanism, there are inherent differences between them that also suggest differences in their effectiveness. According to its proponents (Mitroff, 1982b; Mitroff & Mason, 1981), dialectical inquiry derives from Hegel's dialectic. They argue that high quality decisions will result from pitting diametrically opposed assumptions and recommendations—a thesis and antithesis—against each other. Debate of the two positions that maximizes conflict and critical evaluation will allow a synthesis to evolve that retains assumptions and recommendations that have survived the test. Dialectical inquiry uses two sets of recommendations; devil's advocacy relies on one set and a critique of that set. Mason (1969) suggested that dialectical inquiry should lead to higher quality solutions than devil's advocacy because the latter focuses only on what is wrong with assumptions and recommendations rather than on finding or identifying suitable alternatives. Essentially, devil's advocacy is less effective than dialectical inquiry in developing alternative sets of assumptions and recommendations.

Scholars have also argued that the conflict orientation of both approaches should produce higher quality decisions than group approaches that do not formally structure argumentation and debate (Mason & Mitroff, 1981; Mitroff, Emshoff, & Kilmann, 1979). Schools and organizations commonly teach one such approach, consensus, and top management groups often use it in decision making (Bourgeois, 1980; *Business Week*, 1978; Mintzberg et al., 1976; Quinn, 1980). The consensus approach has been found to lead to higher quality decisions than conventional group interaction does (Nemiroff, Pasmore, & Ford, 1976).[1] Proponents suggest that all members of consensus decision making groups are encouraged to state their assumptions and recommendations and then freely discuss them until they reach final decisions, whereas in conventional groups this is not always the case (Hall, 1971).

Although the consensus approach advocates free expression in groups, it provides no formal procedure for testing and evaluating these expressions. Given the frequent use of consensus in top management groups, it was quite important that we include this approach in the study. Moreover, its inclusion provided an excellent opportunity to compare the effects of formal structured intragroup conflict techniques with a less structured technique.

Although the quality of assumptions and recommendations is important to the effectiveness of decisions by top management groups, group members' satisfaction and desire to continue to work together on subsequent tasks are also important, especially in ongoing groups (Nadler, Hackman, & Lawler, 1979). In structured conflict the inherent argumentation and debate

> can lead to damaged feelings when the members resolutely live up to their role as critical evaluators and take each other's proposals over the bumps. Feelings of rejection, depression, and anger might be evoked so often when this role assignment is put into practice that it could have a corrosive effect on morale and working relations within the group (Janis, 1972: 210).

Results of a survey of American Management Association executives (Schmidt, 1974) demonstrated these concerns. The survey found that although conflict has positive effects, as discussed earlier in this paper, negative outcomes included individuals feeling defeated and demeaned, increased distance between people, suspicion and distrust, and a breaking down of teamwork. In a discussion of decision making in boardrooms, Welty (1983) suggested that consensus creates a less combative environment than debate. Although both dialectical inquiry and devil's advocacy are conflict-based approaches, the first should lead to greater conflict than the second. To the extent that participants have emotional stakes in the assumptions and recommendations they have developed, dialectical inquiry could create greater conflict since it requires the development of diametric positions. Devil's advocacy requires only one set of assumptions and recommendations; those who prepare critiques may develop smaller emotional stakes in their views than do those who develop alternate recommendations. Thus, we offer the following hypotheses concerning the effects of dialectical inquiry, devil's advocacy, and consensus on group performance and individual members' perceptions of groups:

---

[1] It is important to note that the conventional groups used as controls in Nemeroff and colleagues' (1976) study were not instructed as to the process they should use and were left to their own devices in arriving at group decisions. These groups were found to have used majority vote, averaged rankings, trading, and political power among individuals to make decisions. Thus, these groups were not truly control groups. Since groups develop some process, whether instructed to or not, it is much more useful to compare groups trained in particular group approaches than to let them vary.

> *Hypothesis 1: There will be differences in performance between groups using different approaches to group decision making.*
>
> *a: Dialectical inquiry groups will perform better than both devil's advocacy and consensus groups.*
>
> *b: Devil's advocacy groups will perform better than consensus groups.*
>
> *Hypothesis 2: Among groups using different approaches to group decision making, there will be differences in members' satisfaction with groups and in their desires to continue to work with the same groups on subsequent tasks.*
>
> *a: Members of consensus groups will have greater satisfaction and desire to continue working with the groups on subsequent tasks than will members of dialectical inquiry or devil's advocacy groups.*
>
> *b: Members of devil's advocacy groups will have greater satisfaction and desire to continue working with the groups on subsequent tasks than will members of dialectical inquiry groups.*

Since all three group decision making approaches are designed to encourage group members to evaluate and reevaluate critically their assumptions and recommendations, we also attempted to determine which of these methods was most successful in accomplishing this. Therefore, we also tested the following hypotheses:

> *Hypothesis 3: There will be differences in levels of critical evaluation of assumptions and recommendations among members of groups using different approaches to group decision making.*
>
> *a: Dialectical inquiry groups will produce more critical evaluation of assumptions and recommendations among group members than either devil's advocacy or consensus groups.*
>
> *b: Devil's advocacy groups will produce more critical evaluation of assumptions and recommendations among group members than consensus groups.*

METHODS

Subjects

Participants were 120 M.B.A. students at the University of Houston. Their mean age was 29 years (s.d. = 5.2); their mean number of years of full-time work experience was 6.4 (*s.d.* = 4.9); and their mean number of years of managerial experience was 2.3 (*s.d.* = 3.9). With respect to highest level of managerial experience, 45.8 percent had never been managers, 34.2 percent had been lower-level managers, 15.8 percent had been middle-level managers, and 3.3 percent had been top-level managers. Seventy-two were men and 47 were women. Demographic data were missing for one individual.

Subjects were randomly assigned to four-person groups and each group was randomly assigned to either a dialectical inquiry, devil's advocacy, or consensus condition. Thus each experimental condition had ten groups.

Procedures

The study ran simultaneously in three sections of a semester-long M.B.A. course in corporate strategy and policy. We chose this course because the decisions required in it were similar to those for which the three approaches to decision making were designed. The procedures used in the three sections were identical. Participation was a graded course requirement and the study took place during class time. We began collecting data at mid-semester to allow time for the students to have developed the background necessary for making strategic decisions. Students had been exposed to seven weeks of lectures, readings, and cases on strategic management prior to start of the study.

At the end of the class period preceding the actual administration of the experimental conditions, subjects received a copy of the case that would be used in the study. They were instructed to read the case, analyze it, prepare written recommendations and lists of their assumptions, and to be prepared to discuss the case in depth at the next class meeting. To ensure that subjects would be prepared, they were required to turn in copies of their assumptions and recommendations for grading. All complied. It was essential that this initial preparation of the case be conducted at home since it would have otherwise consumed an inordinate amount of class time. It was also reasonable to expect that managers would prepare themselves before group meetings. Subjects were part-time night students; lectures had not previously discussed any of the experimental conditions, and subjects got the case only two days before the experimental session. Thus, it seemed unlikely that preparation at home would confound the research.

Upon the arrival at the experimental session—the next class—subjects learned that they would now be working on the same case in four-person groups. After random assignment to groups they went with escorts to separate rooms for each approach. They then received instructions describing the group approach that they were to use to generate assumptions and recommendations based on the case. Experimenters also read the instructions aloud and answered questions about them. Finally, they isolated groups in their separate locations to begin their work.

To avoid possible experimenter and class section confounds, we took three precautions in designing and administering this experiment. First, groups were equally assigned to experimental conditions across classes.

Second, the three experimenters who gave and read the group instructions were randomly assigned to the different conditions across the three classes. Finally, the experimenters were instructed to answer only questions that requested clarification of the manipulation instructions.

Groups had three hours, amounting to two class sessions, in which to generate and record their final recommendations and assumptions. Subjects were instructed not to discuss the assignment with anyone outside their group or subgroup[2] between class sessions. Each room contained a tape recorder that was used for manipulation checks. All groups were notified when 15 minutes remained in their work time. When time ran out, experimenters returned, collected tape recorders and all materials, and administered a 12-item questionnaire that included manipulation checks and questions eliciting group members' perceptions. A demographic information form was also administered. Subjects were fully debriefed in a subsequent class meeting.

Task

The study used the Leitch Quality Drug Company case (Glueck, 1980). We chose this relatively brief case because it presents subjects with a number of real strategic problems and allows analysis in a reasonable amount of time. The case presents a situation that confronted an actual but disguised company and the four partners who owned it. The company, which operated three drugstores in Orlando, Florida, faced a number of significant changes in its environment as well as a number of internal problems. Some of the changes included demographic shifts and the growth of large super-discount drugstores. The three stores operated in very different neighborhood markets. Internal problems included weak accounting systems, poor inventory control, and ineffective pricing policies. One of the partners, Mr. Neds, had urged the other three to make changes, but to no avail.

Individuals and groups were instructed to analyze the company's situation and address recommendations to Mr. Neds. They were also instructed to state the facts and assumptions that supported their recommendations.

Experimental Manipulations

Dialectical inquiry manipulation. Two studies (Mason, 1969; Mason & Mitroff, 1981) provided the basis for this manipulation. Ten groups received the following statement of information and instruction:

> In the dialectical inquiry approach two different recommendations, based on contrary assumptions, are developed from the same data. The two recommendations and their respective assumptions are subjected to an in-depth, critical evaluation through a debate between two advocate subgroups. Using the same data, the debaters attempt to spell out the implications of each decision, reveal its underlying assumptions, and challenge (or defend) those assumptions as effectively as possible. In other words,

---

[2] Two subgroups were formed with each dialectical inquiry and devil's advocacy group. Each subgroup's responsibilities are defined in the Manipulation section of the paper.

each side is trying to win the debate. Following the debate the two advocate subgroups should settle on which assumptions survived the scrutiny of debate and attempt to develop a recommendation based on them. Proponents of the dialectical inquiry decision making approach believe that sound judgments or recommendations are more likely to result from thorough identification and criticism of proposed decisions and their underlying assumptions. Typical guidelines and procedures used in the dialectical inquiry approach follow:

1. Divide your four-person group into two (2) two-person advocate subgroups.
2a. One subgroup should develop recommendations and build an argument for them, supported by all *key assumptions, facts, and data that underlie them.* Write the recommendations, assumptions, facts, and data as completely and lucidly as you can on the Subgroup I Recommendations form provided for this purpose. Return the form to the other subgroup.
2b. The *other* subgroup should await receipt of the list of key assumptions made by the first group. While waiting, you may discuss the case *only* between yourselves. Upon receiving this list of key assumptions, the second subgroup should develop plausible assumptions which *negate* the first subgroup's key assumptions. Using these new assumptions, the second subgroup should develop counter-recommendations and complete the Subgroup 2 Recommendations form.
3. Following step 2b, both advocate subgroups should present both orally and in writing their assumptions, recommendations, and supporting facts and data to the other subgroup.
4. The two advocate subgroups debate their recommendations and the validity of the assumptions they have made. The objective of this debate is to arrive at a final list of assumptions that is acceptable to both subgroups.
5. Once the debate is completed, you should reach agreement on which assumptions survived. Any new assumptions that arise from the debate should also be included.
6. Using the surviving assumptions, develop recommendations.
7. Record the final recommendations, assumptions, facts, and data on the Final Recommendations form provided.

Devil's advocacy manipulation. This manipulation was based on Cosier (1981) and Mason and Mitroff (1981). Ten groups received the following statement of information and instruction:

The devil's advocacy approach develops a solid argument for a reasonable recommendation, then subjects that recommendation to an in-depth, formal critique. The critique calls into question the assumptions and recommendations presented to the devil's advocate, and attempts to show why the recommendations should not be adopted. Through repeated criticism and revision, the approach leads to mutual acceptance of a recommendation. Proponents of this decision making approach believe that good recommendations and assumptions will survive even the most forceful and effective criticism and that this approach is more likely to yield sound judgments or recommendations. Here are some guidelines and procedures to follow in using the devil's advocate approach:

1. Divide your four-person group into two (2) two-person subgroups. Assign one subgroup the formal role of devil's advocate. Separate into your subgroups.
2. Discuss the Leitch Quality Drugs case with your subgroup partner.
3a. The subgroup that is *not* devil's advocate then should develop a set of recommendations and build an argument for them, supported by all *key assumptions, facts, and data that underlie them.* Write the recommendations, assumptions, facts, and data as clearly and thoroughly as you can on the Subgroup I Recommendations form provided for this purpose.
3b. Meanwhile, the devil's advocate subgroup should prepare for their critique by discussing the case and any critical assumptions, data, etc., which they can identify.

4. The first subgroup presents its written recommendations and assumptions to the devil's advocate subgroup. The devil's advocate subgroup subjects the recommendations to a formal critique. The critique attempts to uncover all that is wrong with the recommendations, assumptions, facts, and data and to expound the reasons why the recommendations should not be adopted.
5. Following Step 4, the critique is presented to the first subgroup orally and on the Critique form provided. The first subgroup then meets separately once again and revises its recommendations to satisfy the valid criticisms of the devil's advocate subgroups.
6. Repeat Steps 4 and 5 until both subgroups can accept the recommendations, assumptions, and data.
7. Write the final recommendations, assumptions, facts, and data on the Final Recommendations form with which you have been provided.

Consensus manipulation. This manipulation was based on Hall (1971) and Nemiroff and colleagues (1976). Ten groups received the following statement of information and instruction:

> The consensus approach relies on a thorough, open, and constructive discussion and examination of the recommendations and underlying assumptions developed individually by the group members. In the course of this discussion each group member should have the opportunity to present his recommendations, the underlying assumptions, and relevant facts and data in as clear and logical a manner as possible. Through discussion, questioning, and more complete exchange of information and opinion, the group seeks a better recommendation than might be produced by a single person. It is not necessary that each person be completely satisfied with the assumptions and recommendations—only that each can accept them on the basis of logic and a willingness to consider them as feasible. *Consensus* is said to exist when *all* group members can accept the assumptions and recommendations on this basis. Proponents of this decision making approach believe that better assumptions and recommendations result from a more complete investigation and airing of data and ideas and a logical resolution of differences within the group. Here are some guidelines to use in achieving consensus:

1. Avoid arguing blindly for your own assumptions and recommendations. Present your position clearly, logically, and persuasively, but consider carefully the comments and reactions of other group members. If you present the same points again, take comments and reactions into account.
2. Avoid making "win-lose" statements in your discussion. Discard the notion that someone must win and someone lose in the discussion. When impasses occur, look for the next most acceptable solution for all parties.
3. Avoid changing your mind simply to avoid conflict and reach agreement. Withstand pressures to yield which have no logically sound foundation. Strive for enlightened flexibility; avoid mere capitulation.
4. Avoid conflict reducing techniques such as majority voting, tossing a coin, and the like. Differences of opinion indicate an incomplete exchange of relevant information on someone's part; press for additional sharing of task or emotional data where it seems in order.
5. View differences of opinion as natural and helpful rather than as a hindrance to decision making. Generally, the more assumptions and recommendations expressed, the greater the likelihood of conflict, and the richer the resources used in solving the problem at hand.
6. View all initial agreements as suspect. Explore the reasons for the apparent agreement. Make sure that people have arrived at similar recommendations either for the same reasons or for complementary reasons before incorporating such recommendations into the group's final set.

7. At the conclusion of the group's work, record the group's final assumptions, recommendations, facts, and data on the Final Recommendations form provided.

Dependent Measures

Group performance. Judges' ratings of the groups' final recommendations and supporting assumptions were used to assess performance. The judges were educators and researchers in business policy and strategy. One of the judges, in fact, wrote the published teaching note for the case used in the study. Two judges independently rated assumptions on three dimensions: (1) the number of assumptions groups brought to the surface, (2) the validity of the assumptions made with respect to information provided in the case, and (3) the importance of the assumptions with respect to the decisions made.

Judges computed scores for numbers of assumptions by counting the assumptions that each group provided. They rated the validity of each assumption on a 5-point scale ranging from 1, low confidence in the validity of the assumption, to 5, high confidence in the validity of the assumption. They rated the importance of each assumption on a 5-point scale ranging from 1, not important, to 5, very important. Each assumption was rated on both its validity and importance. Final group validity scores were determined by summing the validity ratings of each group's assumptions and then dividing the sum by the number of assumptions the group brought to the surface. Judges computed group importance scores in a like manner. The study used validity and importance averages because otherwise scores of groups that surfaced many low-rated assumptions could be inflated beyond those of groups that surfaced fewer, but more key assumptions.

The two judges also independently rated the overall quality of each group's recommendations on a 5-point scale. The scale ranged from 1, low quality recommendation, to 5, high quality recommendation.

Group members' perceptions. A 12-item questionnaire was administered to each group member immediately after a group's work was completed. The questionnaire asked the extent to which respondents agreed with statements describing perceptions of and reactions to the decision making experience. Included in the questionnaire were six items used as manipulation checks, two items for each manipulation; two items assessing satisfaction and desire to continue to work with the groups; two items assessing the extent to which group condition led members to critically reevaluate their personal assumptions and recommendations; and two items assessing acceptance of groups' assumptions and recommendations. The Appendix presents the questionnaire items, response formats, and internal consistency reliabilities.

Tape recordings. Each group's work session was tape recorded for use as a manipulation check. Using as their criteria the three written manipulations provided to the groups, two judges listened to the tapes and independently classified each group into one of the three decision making conditions.

RESULTS

Manipulation Checks

Responses to the six questionnaire items and the tape recordings were used to determine the degree to which experimental manipulations succeeded. The questionnaire items allowed assessment of individuals and the tapes allowed assessment of groups.

Since manipulation checks with two items were used for each condition, we computed internal consistency estimates using Cronbach's alpha. The scales demonstrated reasonable internal consistency (see the Appendix). Therefore, we retained all three checks for further analysis. As expected, agreement with the dialectical inquiry manipulation check differed significantly by condition ($F_{2,117} = 66.7$, $p < .001$). Further, individuals in the dialectical inquiry condition differed significantly from those in the other two conditions ($p < .01$, planned *t*-tests). Agreement with the devil's advocacy manipulation check also differed significantly by condition ($F_{2,117} = 86.1$, $p < .001$). Planned *t*-tests also indicated that individuals in the devil's advocacy condition differed significantly ($p < .001$) from those in the other two conditions. Finally, agreement with the consensus manipulation check also differed significantly by condition ($F_{2,117} = 66.4$, $p < .001$). Planned *t*-tests indicated that individuals in the consensus condition differed significantly from those in the other two conditions ($p < .001$). To further assess the extent to which the manipulations were successful, we conducted a discriminant analysis on the manipulation checks for individuals. The analysis classified 90 percent of all the individual checks correctly as to group condition.

The judges' classifications of the tapes gave less support for the success of the manipulations. The two judges agreed on only 70 percent of the classifications, 21 of 30 classifications made. Of the 9 classifications on which judges did not agree, 7 involved dialectical inquiry or devil's advocacy groups. This suggests that the distinctions between these two approaches are not always apparent in practice. Use of two subgroups in both conditions might have caused overlap. Follow-up discussions between the two judges brought resolution of the disagreements. Although they agreed on final classifications for all 30 tapes, the judges still had questions concerning several of them.

To assess the success of the manipulations, judges' classifications were compared to the actual manipulations. There was a match on 23 of 30 classifications, or 76 percent. Of the seven remaining classifications, three dialectical inquiry manipulations were classified as devil's advocacy, two devil's advocacy manipulations were classified as dialectical inquiry, one dialectical inquiry was classified as consensus, and one consensus was classified as dialectical inquiry. Overall, the individual level manipulation checks suggest that the manipulations were successful. However, the taped manipulation checks provided only moderate support for the success of the manipulations. The discrepancy between the individual and taped manipulation checks may be partially due to the priming effects that the manipulation instructions may have had on individuals' questionnaire responses.

Group Performance[3]

Table 1 presents means and standard deviations for the measures of group performance. Table 2 summarizes the results for each measure, which are also discussed in the next section.

Number of assumptions. Before applying the number of assumptions measure, we assessed interrater reliability using correlations and percentage of agreement between judges. The correlation between the judges was .74; they agreed on 76 percent or on 23 of 30 groups. In only one case was there large disagreement between the judges; this disagreement focused on whether certain statements were in fact assumptions. Since six of the differences were small, we resolved them through averaging; the seventh was resolved through discussion. The number of assumptions groups brought to the surface did not differ significantly across the experimental conditions.

Quality of assumptions. Two measures were used to assess the quality of the assumptions groups surfaced: validity and importance. For both measures of quality, we computed interrater correlations and percentage of agreement

TABLE 1
Means and Standard Deviations for Group Performance and Group Members' Perceptions

	Type of Group Approach					
	Dialectical Inquiry		Devil's Advocacy		Consensus	
Dependent Variables	Means	s.d.	Means	s.d.	Means	s.d.
Group performance						
Number of assumptions	12.50	3.57	11.65	5.96	13.40	5.47
Validity of assumptions	4.27	0.39	3.81	0.52	3.16	0.30
Importance of assumptions	4.02	0.29	3.72	0.50	3.06	0.23
Group members' perceptions						
Satisfaction with and desire to continue to work with the group	7.23	1.46	7.13	1.24	8.23	1.41
Critical evaluation	8.50	1.11	8.38	1.23	7.03	1.37
Acceptance of group decision	6.50	1.20	7.30	1.36	8.15	1.00
Confidence in group decision	7.23	1.33	7.60	1.32	7.68	0.89

---

[3] Because of the questionable results of the taped manipulation checks, we conducted analyses parallel to the ones presented in this section without the seven groups for which the judges' classifications did not agree with the actual manipulations. All analyses were identical for the full sample of 30 groups and the sample excluding the seven groups, except for the analysis involving critical evaluation. It was not significant in the smaller sample.

between judges. We eliminated redundant assumptions among the groups and assessed reliabilities on all unique assumptions. A total of 297 unique assumptions were judged. For the validity measure the correlation was .88. The actual level of agreement was 66 percent, with 22 percent of the assumption ratings deviating by one point on the 5-point scale and 9 percent deviating by two points. The correlation for the importance measure was .78. The actual level of agreement was 59 percent with 28 percent of the assumption ratings deviating by one point on the 5-point scale and 5 percent deviating by two points. Given the extreme difficulty in accurately rating so many assumptions, we concluded that a reasonable level of interrater reliability had been established to retain both measures. Because differences between judges were not very large, we averaged the judges' ratings to arrive at the final rating.

The correlation between the validity and importance measures was .77 ($p<.001$). This correlation and conceptual evidence suggesting that these measures are related required that we analyze validity and importance together; the combination was termed *quality of assumptions.* A one-way MANOVA indicated that there was a significant difference in the quality of assumptions that surfaced among the conditions ($F_{4,52} = 9.06$, Wilks's lambda = .347, $p<.001$). To further understand the univariate effects, we computed F-tests for each dependent variable. The effects for both validity ($F_{2,27} = 17.88$, $p<.001$, $\omega^2 = .53$) and importance ($F_{2,27} = 18.4$, $p<.001$, $\omega^2 = .54$) were significant. As indicated in Table 2, all but one of the planned t-tests among the groups were also significant for both validity and importance of assumptions. The exception was the comparison between dialectical inquiry and devil's advocacy for importance of assumptions.

Quality of recommendations. Using the ratings for all 30 groups, we computed both correlation and percentage of agreement again for the two judges' ratings. In this case the correlation between the judges was .68. The level of agreement was 60 percent with 10 percent of the quality ratings deviating by one point on the 5-point scale and 10 percent deviating by two points. Twenty percent of the ratings deviated by three or more points. The moderate level of disagreement suggested that discussion rather than averaging would be the best way to resolve differences between the judges. All discrepancies were resolved. Given the moderate correlation and percentage of agreement between the judges, some caution is required in interpreting the results using this measure.

The results on decision quality indicate that there were significant differences among groups ($F_{2,27} = 3.73$, $p<.05$, $\omega^2 = .08$). Further analysis using planned t-tests (see Table 2) found that both the dialectical inquiry and devil's advocacy groups made significantly higher quality recommendations than the consensus groups, but that there was no difference in quality between the first two types.

Of additional interest was the relationship between the quality of recommendations and the quality and number of assumptions brought to the surface. The correlation between quality of recommendations and number of assumptions was not significant ($r=.10$). The correlation between the quality of

TABLE 2
Results of Performance Comparisons among Group Conditions[a]

Dependent Measures	Results
Quality of assumptions	
MANOVA	
Validity and importance	$F_{4,52}$ = 9.06***
ANOVA	
Validity	$F_{2,27}$[b] = 17.88***
Importance	$F_{2,27}$ = 18.43***
Number of assumptions	$F_{2,27}$ = 1.33
t-tests	
Validity	
Dialectical inquiry vs. devil's advocacy	t_{18} = 2.23*
Dialectical inquiry vs. consensus	t_{18} = 7.12***
Devil's advocacy vs. consensus	t_{18} = 3.37**
Importance	
Dialectical inquiry vs. devil's advocacy	t_{18} = 1.53
Dialectical inquiry vs. consensus	t_{18} = 7.18***
Devil's advocacy vs. consensus	t_{18} = 3.77**
Quality of recommendations	
ANOVA	
Quality of recommendation	$F_{2,27}$ = 3.73*
t-tests	
Quality of recommendations	
Dialectical inquiry vs. devil's advocacy	t_{18} = 1.23
Dialectical inquiry vs. consensus	t_{18} = 2.34*
Devil's advocacy vs. consensus	t_{18} = 2.49*

[a] N = 30, 10 per cell.

[b] Subscripts are degrees of freedom.

$^{*}p<.05$

$^{**}p<.01$

$^{***}p<.001$

recommendations and the quality of assumptions—the combined validity and importance scores—was barely significant ($r = .39, p<.10$). Although not very strong, the latter correlation supports numerous decision theorists who have suggested that bringing assumptions to the surface and analyzing them are crucial for high quality recommendations. We can make no causal inferferences here about how individuals thought through the assumptions and recommendations, since it is possible that they first made their recommendations and only afterwards supported them with assumptions. This possibility was considered unlikely, but it should be noted that this study had no controls to prevent it. With respect to the groups themselves, the contents of the tape recordings indicate that all the groups initially arrived at their assumptions prior to developing the recommendations. However, a few groups modified assumptions as they evaluated recommendations; there was no relationship between experimental condition and these modifications.

Group members' perceptions. Table 1 presents means and standard deviations for the measures of group members' perceptions. Table 3 presents results for each individual measure, which are also discussed next in this section.

Satisfaction and desire to continue to work with groups. We used a 2-item scale developed from Tjosvold and Field (1983) to measure satisfaction with groups. Since the scale demonstrated reasonable internal consistency we retained it for further analysis. In addition to performance differences among the three conditions, there were also significant differences in satisfaction and desire to work with groups ($F_{2,117} = 7.86$, $p<.001$, $\omega^2 = .10$). Planned *t*-tests presented in Table 3 indicate that individuals in the consensus groups were significantly more satisfied and desirous of working with their groups in the future than members of either the devil's advocacy or dialectical inquiry groups. No difference emerged, however, between the latter two types of group. These data suggest that programmed conflict techniques lead to lower satisfaction among group members.

Critical evaluation. The 2-item measure used to assess this construct demonstrated low internal consistency. Therefore, caution is needed in interpreting the findings using it. The results indicate a significant difference among the conflict methods in their ability to induce members to subject their own ideas to critical evaluation ($F_{2,117} = 17.36$, $p<.001$, $\omega^2 = .04$). Planned *t*-tests (see Table 3) indicated that both dialectical inquiry and devil's

TABLE 3

Group Members' Perceptions across Group Conditions[a]

Dependent Measures	Results
Satisfaction and desire to continue to work with group	
ANOVA	$F_{2,117}$[b] = 7.86***
t-tests	
Dialectical inquiry vs. devil's advocacy	t_{78} = -3.71***
Dialectical inquiry vs. consensus	t_{78} = -3.12**
Devil's advocacy vs. consensus	t_{78} = -0.33
Critical evaluation of assumptions	
ANOVA	$F_{2,117}$ = 17.36**
t-tests	
Dialectical inquiry vs. devil's advocacy	t_{78} = 0.48
Dialectical inquiry vs. consensus	t_{78} = 5.30***
Dialectical advocacy vs. consensus	t_{78} = 4.63***
Acceptance of group decisions	
ANOVA	$F_{2,117}$ = 19.02***
Duncan Multiple Range test	
Dialectical inquiry vs. devil's advocacy	W > .688**
Dialectical inquiry vs. consensus	W > .718**
Devil's advocacy vs. consensus	W > .688**

[a] N=120, 40 per cell.

[b] Subscripts are degrees of freedom.

*p<.05

**p<.01

***p<.001

advocacy were more capable of stimulating critical evaluation than consensus, but that there was no difference between dialectical inquiry and devil's advocacy.

Acceptance of group decisions. Although no hypothesis addressed it, we took an additional measure commonly used in research on group decision making (Tjosvold & Field, 1983), acceptance of a group decision. A 2-item acceptance measure yielded reasonable internal consistency.

There was a significant difference among the group conditions on acceptance ($F_{2,117} = 19.02, p<.001, \omega^2 = .24$). Since differences were not hypothesized a priori, specific differences among the group means could not be assessed with planned comparisons. However, post hoc comparisons using Duncan's multiple range test (see Table 3) indicated that there was greater acceptance of decisions with the consensus condition than with either of the two other conditions. The devil's advocacy condition also led to greater acceptance than the dialectical inquiry condition.

DISCUSSION

This study, the first controlled laboratory experiment to assess the comparative effectiveness of dialectical inquiry, devil's advocacy, and consensus in a group context, suggests that programmed conflict is useful in improving the quality of strategic decisions. In particular, the results obtained using both the full set of groups and excluding the seven groups suggest that both dialectical inquiry and devil's advocacy were more effective than consensus in generating high quality recommendations and underlying assumptions. Although dialectical inquiry was more effective than devil's advocacy with respect to the quality of the assumptions, there were no differences with respect to the quality of the recommendations. These results support Mitroff's general proposition that dialectical inquiry is more effective than devil's advocacy in bringing high quality assumptions to the surface, but fails to confirm the proposition that dialectical inquiry is more effective than devil's advocacy in the development of quality recommendations. Although the previously cited studies by Cosier and his colleagues have been criticized on conceptual grounds (Mitroff, 1982b; Schweiger & Finger, 1984), this study provided partial support for their findings. That is, it did not show dialectical inquiry to be clearly superior to devil's advocacy. Since there were significant differences between the two approaches in the assumptions generated and not in the recommendations, further research is suggested on how these two approaches function.

Very little is known about how groups should actually use either approach to arrive at decisions. All that theorists have suggested is that a debate follow the critique in devil's advocacy and the negation in dialectical inquiry. Since neither method clearly specifies how groups should use acceptable sets of assumptions to construct final recommendations, it is conceivable that any benefits accrued in bringing assumptions to the surface are diluted as recommendations are developed. More research is needed to understand this transition from assumptions to recommendations. Although the correlation was not very strong, the quality of recommendations was significantly related to the quality of assumptions. Again, these findings support the views of Mitroff and his colleagues. The lowness of this correlation, however, suggests that the assumptions and some degree of error were not the only factors affecting the quality of recommendations. Understanding the processes by which groups develop assumptions and use them to choose recommendations may be quite instrumental to understanding the limits of these decision making approaches. It is possible that the processes used to arrive at assumptions and recommendations differ. The strength of the ω^2s further suggests that the group approaches had more of an effect on the quality of assumptions raised than on the quality of recommendations. It is also possible that these approaches are easier to describe than to carry out. The taped manipulation checks may indicate that group members need to be trained in using the approaches or need to gain experience with them before they can successfully implement them.

The results also suggest that those interested in selecting group decision making techniques should pay attention to context. Although consensus may be useful for some nonstrategic tasks, the evidence provided in this study clearly suggests that it is inferior to both dialectical inquiry and devil's advocacy with respect to the overall quality of the assumptions and recommendations made. This is not surprising, given Janis's (1972) concern that consensus may lead to too little conflict in the group decision making process and thus to uncritical acceptance of invalid assumptions and inferior recommendations. The results of this study partially support the idea that both dialectical inquiry and devil's advocacy lead to significantly greater critical evaluation of personal assumptions on the part of group members than consensus. However, problems of measurement and the nonsignificant results obtained from the analysis excluding the seven groups preclude any firm conclusions. More research is certainly warranted on the application of these techniques to specific types of problems and decisions.

Concerning satisfaction with groups and desire to work with them in the future, the findings suggest that consensus may be more functional than either dialectical inquiry or devil's advocacy for preserving harmony within a group. This was also evident in the greater degree of acceptance of decisions in the consensus groups than in the dialectical inquiry and devil's advocacy groups. Overall, these results pose a dilemma in selecting a decision making method. On the one hand, conflict improves decision quality; on the other, it may weaken the ability of a group to work together in the future. Obviously, this is less of a problem for ad hoc groups than for ongoing ones like top management teams. It is plausible that an ongoing group's ability to continue using the conflict procedures might decline over time. This research provided no evidence that explained why members of the dialectical inquiry and devil's advocacy groups were less satisfied. Perhaps people just do not view conflict favorably. A more likely explanation, however, is that the groups did not manage conflict properly. An actual organizational group needs norms and rewards that encourage constructive conflict (Brodwin & Bourgeois, 1984; Sussman & Herden, 1982).

The findings on satisfaction and acceptance also have implications for decision implementation. Group members may be more likely to implement decisions that were reached through low rather than high conflict approaches. Again, this presents a dilemma—group members may not accept high quality decisions. Since we did not monitor and analyze the actual group processes involved in the three conditions, other than through the taped manipulation checks, it is difficult to determine why conflict reduced satisfaction and acceptance. Unfortunately, the difficulty of identifying individual interactions in our tapes prevented such an analysis. Future research needs to more fully address process in the three approaches.

As does any study, this one has its limitations. Although a strategic decision making case was the task, it had certain weaknesses. Since the company was privately held, subjects received all available information from one source. In most strategic decisions, however, decision makers have to determine what information is needed and where to find it. In addition, those actually involved in the company would already have had certain industry, company, and task group experiences that subjects in this study did not begin with.

A second limitation was the population of subjects used. These were not real-life strategic decision makers. However, the demographic data suggested that subjects had a reasonable degree of managerial experience. Also, subjects had studied a reasonable number of strategic concepts prior to working on the case. Even so, the task and the subjects suggest some caution in generalizing these findings to actual organizational strategic decisions and decision makers.

A third limitation is the time frame of the study. The experiment ran for one week. Although managers may make many decisions this quickly, it is unlikely that comprehensive strategic decisions, such as were required here, would have been made in such a short time. Thus, the time frame may have affected the results.

Fourth, the study did not include time as a performance variable. Groups had a fixed period within which to complete the task and had little incentive to finish more quickly. In practice, however, top management's time carries a high opportunity cost. Moreover, speed is a crucial consideration in some decisions. The best choice of decision making approach may depend on the expected value of incremental decision quality, the cost of additional management work hours devoted to the decision, and the urgency of the situation.

Finally, this study, like most group decision making studies conducted in laboratory settings, relies exclusively on ad hoc groups. Clearly, longitudinal studies of ongoing groups are needed. In particular, such studies may provide additional insight into the long-term effects of programmed group conflict on groups. One potential effect this study identified is the decline in group members' willingness to work together after experiencing the conflict induced by dialectical inquiry or devil's advocacy. A second contribution of longitudinal studies could be to identify effects of learning on both individuals and groups brought about by repeated exposure to these group decision approaches. Such learning could influence the relative effectiveness of these approaches over time. A final contribution of longitudinal studies would be to examine the effects that development of group norms and of shared information, perceptions, and values among members might have on sustaining the use of group conflict approaches (Beyer, 1981).

REFERENCES

Beyer, J. M. 1981. Ideologies, values, and decision making in organizations. In P. Nystrom & W. H. Starbuck (Eds.), ***Handbook of organizational design***, vol. 2: 166–197. London: Oxford University Press.

Bourgeois, L. J. 1980. Performance and consensus. ***Strategic Management Journal***, 1: 227–248.

Brodwin, D. R., & Bourgeois, L. J. 1984. Five steps to strategic action. In G. Carroll & D. Vogel (Eds.), ***Strategy and organization: A West Coast perspective***: 167–181. Boston: Pitman.

Business Week. 1978. Olin's shift to strategy planning. March 27: 102–105.

Cosier, R. A. 1978. The effects of three potential aids for making strategic decisions on prediction accuracy. ***Organizational Behavior and Human Performance***, 22: 295–306.

Cosier, R. A. 1980. Inquiry method, goal difficulty, and context effects on performance. ***Decision Sciences***, 11: 1–16.

Cosier, R. A. 1981. Dialectical inquiry in strategic planning: A case of premature acceptance? ***Academy of Management Review***, 6: 643–648.

Cosier, R. A. 1982. Methods for improving the strategic decision: Dialectic versus the devil's advocate. ***Strategic Management Journal***, 3: 373–374.

Cosier, R. A. 1983. Approaches for the experimental examination of the dialectic. ***Strategic Management Journal***, 4: 79–84.

Cosier, R. A., & Aplin, J. C. 1980. A critical view of dialectical inquiry as a tool in strategic planning. ***Strategic Management Journal***, 1: 343–356.

Cosier, R. A., Ruble, T. A., & Aplin, J. C. 1978. An evaluation of the effectiveness of dialectical inquiry systems. ***Management Science***, 24: 1483–1490.

Glueck, W. F. 1980. ***Business policy and strategic management*** (3rd ed.). New York: McGraw-Hill Book Co.

Hall, J. 1971. Decisions, decisions, decisions. ***Psychology Today***, 5: 51–54, 86, 88.

Janis, I. L. 1972. ***Victims of groupthink: Psychological studies of foreign policy decisions and fiascoes***. Boston: Houghton-Mifflin.

Lourenco, S. V., & Glidewell, J. C. 1974. A dialectical analysis of organizational conflict. ***Administrative Science Quarterly***, 20: 489–508.

Mason, R. O. 1969. A dialectical approach to strategic planning. ***Management Science***, 15: B403–B414.

Mason, R. O., & Mitroff, I. I. 1981. ***Challenging strategic planning assumptions***. New York: Wiley-Interscience.

Mintzberg, H., Raisinghani, D., & Theoret, A. 1976. The structure of unstructured decision processes. ***Administrative Science Quarterly***, 21: 246–275.

Mitroff, I. I. 1982a. Talking past one's colleagues in matters of policy. ***Strategic Management Journal***, 3: 374–375.

Mitroff, I. I. 1982b. Dialectic squared: A fundamental difference in perception of the meanings of some key concepts in social science. ***Decision Sciences***, 13: 222–224.

Mitroff, I. I., & Emshoff, J. R. 1979. On strategic assumption-making: A dialectical approach to policy and planning. ***Academy of Management Review***, 4: 1–12.

Mitroff, I. I., & Mason, R. O. 1981. The metaphysics of policy and planning: A reply to Cosier. ***Academy of Management Review***, 6: 649–651.

Mitroff, I. I., Barabba, V. P., & Kilmann, R. H. 1977. The application of behavioral and philosophical technologies to strategic planning: A case study of a large federal agency. ***Management Science***, 24: 44–58.

Mitroff, I. I., Emshoff, J. R., & Kilmann, R. H. 1979. Assumptional analysis: A methodology for strategic problem solving. ***Management Science***, 25: 583–593.

Nadler, D. A., Hackman, J. R., & Lawler, E. E. 1979. ***Managing organizational behavior***. Boston: Little, Brown & Co.

Nemiroff, P. M., Pasmore, W. A., & Ford, D. L. 1976. The effects of two normative structural interventions on established and ad hoc groups: Implications for improving decision making effectiveness. ***Decision Sciences***, 7: 841–855.

Quinn, J. B. 1980. ***Strategies for change: Logical incrementalism***. Homewood, Ill.: Richard B. Irwin.

Schmidt, W. H. 1974. Conflict: A powerful process for (good or bad) change. ***Management Review***. 63: 4–10.

Schweiger, D. M., & Finger, P. A. 1984. The comparative effectiveness of Dialectical Inquiry and Devil's Advocacy: The impact of task biases on previous research findings. ***Strategic Management Journal***, 5: 335–350.

Schwenk, C. R. 1982a. Why sacrifice rigour for relevance? A proposal for combining laboratory and field research in strategic management. ***Strategic Management Journal***, 3: 213–225.

Schwenk, C. R. 1982b. Effects of inquiry methods and ambiguity tolerance on prediction performance. ***Decision Sciences***, 13: 207–221.

Schwenk, C. R. 1982c. Dialectical inquiry in strategic decision making: A comment on the continuing debate. ***Strategic Management Journal***, 3: 371–373.

Schwenk, C. R. 1983. Laboratory research on ill-structured decision aids: The case of dialectical inquiry. ***Decision Sciences***, 14: 140–144.

Schwenk, C. R. 1984. Effects of planning aids and presentation media on performance and affective responses in strategic decision making. ***Management Science***, 30: 263–272.

Schwenk, C. R., & Cosier, R. A. 1980. Effects of the expert, devil's advocate, and dialectical inquiry methods on prediction performance. ***Organizational Behavior and Human Performance***, 26: 409–424.

Springer, C., & Hofer, C. 1978. General Electric's evolving management system. In C. W. Hofer, E. A. Murray, R. Charan, & R. A. Pitts (Eds.), ***Strategic management: A casebook in business policy and planning***: 454–470. St. Paul, Minn.: West.

Stagner, R. 1969. Corporate decision making. ***Journal of Applied Psychology***, 53: 1–13.

Sussman, L., & Herden, R. 1982. Dialectical problem solving. ***Business Horizons***, 25: 66–71.

Tjosvold, D., & Field, R.H.G. 1983. Effects of social context on consensus and majority vote decision making. ***Academy of Management Journal***, 26: 500–506.

Welty, J. D. 1983. Debate vs. consensus in the boardroom. ***Directors and Boards***, 8: 34–37.

APPENDIX

Manipulation Checks on Individuals and Items from Group Member Perception Questionnaire

Six questionnaire items (two per manipulation) were used as manipulation checks. Group members evaluated the degree to which each item described what their group had done in arriving at its assumptions and recommendations. Responses were on the following 5-point scale: 1, strongly disagree; 2, disagree; 3, undecided; 4, agree; 5, strongly agree.

Two items measured the dialectical inquiry manipulation: (1) the two subgroups considered two conflicting sets of recommendations based on opposite sets of assumptions; and (2) on the basis of a debate over two sets of opposing assumptions developed by each subgroup, the group developed a pool of valid assumptions and based its recommendations on these assumptions.

Two items measured the devil's advocacy manipulation: (1) one subgroup initially developed a set of recommendations followed by a formal critique of it by a second subgroup; and (2) the group's final set of recommendations was based on the outcomes of a formal critique by one subgroup of the initial recommendations of another subgroup.

Two items measured the consensus manipulation: (1) during the group meeting only, members of the group individually presented their personal recommendations and assumptions; and (2) the group's final set of recommendations was the combined, but not necessarily the averaged outcome of members' individual contributions.

Alphas for the two-item dialectical inquiry, devil's advocacy, and consensus manipulation checks were .88, .84, and .63, respectively.

Six additional items measured participants' reactions to their group experience. Group members evaluated the degree to which each statement described their feeling on the following 5-point scale: 1, strongly disagree, 2, disagree; 3, undecided; 4, agree; 5, strongly agree.

Two items measured satisfaction and desire to continue to work with the groups: (1) I would be willing to work with this group on other projects in the future; (2) working with my group was an enjoyable experience.

Two items measured the extent to which the group condition led members to critically reevaluate their own assumptions and recommendations: (1) the group decision process made me critically reevaluate the validity of the assumptions and recommendations that I made in my individual analysis; (2) the group decision process uncovered valid recommendations and assumptions that I did not consider in my individual analysis.

Two items measured group decision acceptance: (1) I am committed to my group's recommendations and assumptions; (2) I am satisfied with my group's recommendations and assumptions.

Alphas for the satisfaction, critical reevaluation, and acceptance measures were .80, .79, and .48, respectively.

David M. Schweiger earned his D.B.A. degree at the University of Maryland. He is Assistant Professor of Strategic Management at the University of Houston. His current research interests include strategic decision making, the effects of mergers and acquisitions on human resources, and venture capitalist decision making.

William R. Sandberg is Assistant Professor of Strategic Management at the University of Houston. He received his Ph.D. degree from the University of Georgia. His research interests include new venture strategies, industry structure, and strategic decision making.

James W. Ragan is a Ph.D. candidate at the University of Houston and received his J.D. degree from the University of South Carolina. Besides managerial decision making interventions, his research interests include work place discipline and the effects of governmental regulation on individuals and business organizations.

© *Academy of Management Journal*
1986, Vol. 29, No. 1, 72–83.

ADMINISTRATIVE SUCCESSION AND ORGANIZATIONAL PERFORMANCE: HOW ADMINISTRATOR EXPERIENCE MEDIATES THE SUCCESSION EFFECT

JEFFREY PFEFFER
ALISON DAVIS-BLAKE
Stanford University

Analysis of data for 22 teams in the National Basketball Association (NBA) indicated that succession has no effect on subsequent team performance when prior performance is controlled, a result consistent with those of other studies. However, when the competence of new coaches is included in the analysis, it appears that succession affects subsequent performance. Teams in which succession occurred, and in which (1) the coach had a good prior record, (2) had prior experience in the NBA or ABA, or (3) had brought about performance improvements in other teams, performed better than teams whose new coaches either lacked experience or had performed less well in the past. The results suggest that attempts to estimate succession effects need to consider the competence of successors.

The effect of managerial replacement on organizational performance remains a debated issue in spite of a number of studies and renewed attention to the problem (e.g., Allen, Panian, & Lotz, 1979; Brown, 1982). As Allen and his coauthors noted, there are "at least three theories about the relationship of managerial succession to organizational performance" (1979: 167). One theory claims that succession improves performance; a second asserts that succession disrupts performance; and a third maintains that succession has no effect on performance. Studies of the effects of managerial succession on performance have evolved over time. Recent studies have controlled for past performance and for effects of regression to the mean. They have considered the differential effects of timing of succession—whether it occurs between seasons or during a season in the case of sports teams. They have also considered the differential effects of inside or outside succession (Allen et al., 1979).

Research studying succession in organizations has had two major theoretical thrusts. First, there has been interest in the effect of succession per se, in such questions as whether turnover negatively affects performance (Grusky, 1963) and in the symbolic and ritualized meanings and forms of succession (Gephart, 1978). Second, scholars have assumed that studying succession will generate information on the vexing issue of whether leaders or managers make a difference (Pfeffer, 1977; Smith, Carson, & Alexander, 1984). In one such study, Lieberson and O'Conner (1972) used succession events as variables in partitioning variation in profits, sales, and profit margins to the sources of organization, time, and leadership. Salancik and Pfeffer (1977) looked at variations in budget allocations in cities in a similarly designed study.

Prior to the recent study by Smith and his coauthors (1984), no analysis of succession events had addressed the experience or abilities of successors. However, examining successors' abilities is necessary to fully explore both of the theoretical problems that have prompted research on succession. Many studies of the effects of succession have found that the event either had no effect or a very small effect on performance (Allen et al., 1979; Brown, 1982; Eitzen & Yetman, 1972). Explanations for lack of effects have often hinged on the abilities or experience of the successors. For example, March and March (1977) argued that there is often a filtering process in organizations that leads to selection from a fairly homogeneous group. If the individuals selected are fairly similar to those they succeed in terms of ability, strategy, and style, the effect of succession will be slight. Similarly, organizations that are performing poorly and are therefore more likely to experience succession (Allen et al., 1979; Salancik & Pfeffer, 1980) may not be able to attract the best replacements. In a competitive market for managerial talent, people who perform well are likely to be interested in going to organizations that do the same. Unless analysis explicitly includes successors' abilities, absence of a succession effect may be due to similarity in the abilities of the different managers or to a succession's having, in fact, no effect on performance.

Similarly, when succession events are used to examine the effect of leaders on performance, it seems clear that who succeeds becomes as important as the nature of a succession event itself. Therefore, investigators should consider both the characteristics and timing of succession events and the backgrounds and abilities of successors.

The purpose of this study was to examine the effect of successors' abilities on the results of succession in 22 National Basketball Association (NBA) teams. In the course of the analysis we also replicated previous findings of no succession effect for succession considered by itself.

BACKGROUND

There is little dispute that one cause of succession is poor organizational performance (Allen et al., 1979; Allen & Panian, 1982; Brown, 1982; Salancik & Pfeffer, 1980). What remains in question is the effect of succession on subsequent organizational performance. The early succession literature viewed succession as a disruptive event that depressed organizational performance. Gouldner (1954) observed that a change of managers in a gypsum plant disrupted operations and ultimately led to a wildcat strike. Grusky (1963, 1964) developed a vicious-circle theory of succession based on his study of baseball teams. He argued that teams are more likely to experience turnover when they are performing poorly, but that once succession occurs its disruptive effects will tend to cause further deterioration in performance.

The common sense view is that changing managers responsible for poor performance should make for better performance in the future. Although there is some case study data supporting the positive effects of succession (Guest, 1962), very little other empirical support exists. Rather, the most common finding in the literature is that succession has no effect on performance. In a study of college basketball teams, Eitzen and Yetman (1972) found that once they controlled for previous levels of team success, a change in coaches made no difference in performance. Allen and his colleagues found a succession effect on the performance of major league baseball teams, and further noted that "succession between seasons is more likely to improve performance than succession during a season . . . inside succession is less likely to disrupt team performance than outside succession" (1979: 178). However, the size of the effects were small, so that they concluded that "a change in managers typically has only a small impact on subsequent performance" (1979: 179). Brown also concluded that "there is little in the way of a succession effect among professional football organizations" (1982: 13). In a study of succession in a sample of Methodist churches, Smith, Carson, and Alexander (1984: 772) reported no succession effect for the dependent variables of attendance, membership, property value, total giving, or operating budget; this result held for both one-year and three-year lags.

There are several possible explanations for findings of no succession effects or only small effects. One possibility is that two simultaneous forces—a positive effect from the replacement of an unsuccessful manager and a negative effect resulting from the disruption occasioned by the succession—may cancel each other out. Another possibility is that succession is an example of ritual scapegoating (Gamson & Scotch, 1964) and that for a variety of reasons managers do not matter much (Pfeffer, 1977). A third possibility is that managers do matter (Weiner & Mahoney, 1981), but that in order to observe a managerial effect, investigators need to consider the characteristics or abilities of managers. Much of the time the kind of filtering process discussed by March and March (1977) and Pfeffer (1977) may work to insure that new managers are similar to those being replaced. In such circumstances, little if any succession effect is likely to occur. Replacement could be expected to have an effect only in those instances in which new managers differ from the old ones. Moreover, the direction of the effect should be related to successors' abilities.

Smith and his colleagues (1984) investigated the effects of managerial ability on organizational performance following succession. Several characteristics of their study suggest that further investigation of this issue would be useful. They studied the careers of 50 Methodist ministers, computing a salary equation regressing current salary on years of service in the ministry for each individual. Ministers who received salaries more than one standard deviation above the mean eight or more times were labeled as effective; 7 of the 50 were so identified. Then, they conducted an analysis regressing various performance indicators such as attendance and giving in ministers' last years with churches on the values of those indicators in the ministers' first years of service; a zero-one variable indicated whether or not ministers were high-performing. They inferred the effect of leadership ability on performance from the statistically significant effects of this dummy variable in the regressions.

The study (Smith et al., 1984) raised many issues. First, ministers, not churches, were the units of analysis. Instead of following a given church over time and examining the effects of performance of ministers of varying ability, Smith and colleagues sampled ministers and followed them across churches. Assignment of ministers to churches is unlikely to be random, and the analysis reported in the article did not allow separation of the effects of the assignment process from the effects of the succession event itself. Also, the analysis did not test for effects of regression to the mean. Moreover, is salary the best predictor of ability, particularly in a labor market that is not likely to follow many of the dictates of perfect competition? This is especially problematic because the equation used to estimate the ministers' effectiveness did not control for congregational size or resources and thus possibly confounded these variables with effectiveness. Conceptually, it would have been possible to estimate leaders' performance levels from their past success in improving each of the dimensions of organizational performance and then to use this estimate as a subsequent control or indicator of leader effectiveness. Finally, the use of only extreme values of leader effectiveness seems to limit the explanatory power of the study. If leader ability matters, it is probably a more or less continuous variable and might be measured as such instead of with a zero-one value. These limitations of Smith and his colleagues' study and the importance of the leadership/succession issue suggest a need for further empirical research.

The current research thus both replicates and extends past research on succession. Most past research has tended to treat all managers as equal and has ignored variation in characteristics that might be related to their ability to affect organizational performance. It seems reasonable to further investigate the possibility that all managers—or, in the case of sports teams, all coaches—do not have equal ability and that those with superior ability might be better able to generate improvement in performance. Failure to account for the ability of a successor confounds that ability with the succession event itself. The present research both examines the effects of succession and investigates the effects of variation in ability on the consequences of succession. Further, this study attempted to overcome some of the limitations of Smith and colleagues' study by examining several organizations over time and by directly measuring past performance of new managers.

DATA AND METHODS

This analysis used pooled time series and cross-sectional data for all of the teams in the NBA for the 1977 through 1981 seasons. The sample included no teams that entered the NBA after 1977. Twenty-two teams were in the NBA for all of the five seasons, resulting in a total of 110 observations. Succession within or between seasons occurred in 27 of the 110 instances studied. We obtained data from the *Sporting News Official NBA Guide* for each of the seasons included in the sample. Analyzing pooled cross-sectional and time series data involves problems with autocorrelation that make the interpretation of ordinary least squares regression difficult (Hannan & Young, 1977). Therefore, most of the analyses reported here derive from generalized least-squares (GLS) techniques. The particular GLS procedure employed in this analysis is difficult to use with unbalanced designs—for example, without the same numbers of observations or cases for each year. This restricted the analysis to these five years if we wanted to use the 22 teams. Alternatively, we could have used a smaller sample extending back further in time, but there is no reason to expect that there would have been any difference in the results.

Another issue is our choosing to study basketball teams. Several factors motivated the choice of this particular type of organization.[1] First, examining the effects of succession and the abilities of successors demands a setting in which succession and organizational and managerial performance are both readily measured. As Lenz (1981) noted, there are a number of problems and difficulties in studying the performance of business firms. Also, selecting a measure of managerial performance is extremely difficult because it is not clear what the measure should be and because many performance measures are either distorted or unavailable. In sports teams, performance and succession measures are readily available and relatively easy to interpret. Although future studies of succession should be conducted in settings in which performance is less easily measured, these studies will need to carefully consider selection of appropriate performance measures. In particular, these studies will need to insure that measures of managerial performance are responsive to managerial effort and are not determined primarily by forces outside managerial control. Second, because this research attempted to extend an idea developed in the study of sports teams, sports teams seemed to be the most appropriate units on which to test our hypotheses. If a different type of organization were used and the extension of the original ideas was not confirmed, it would be difficult to determine whether our hypotheses or differences in the types of organizations in which they were tested were responsible. Attempting to extend the ideas in their original context eliminates this potential problem in interpreting the results.

As Brown (1982) noted, baseball is very much an individual sport in which strategy and coaching probably have only a slight effect on performance. Although football teams are, in many respects, more like small business organizations than baseball teams are, their larger coaching staffs and player rosters and shorter seasons mean that random variations in performance can have significant effects on win-loss percentages. Basketball, a team sport with essentially one coach and many games played during a season, appears to be a useful starting point for the analysis. Thus, we chose to study the 22 professional basketball teams in the NBA.

---

[1] Clearly the results reported here require replication and generalization in other contexts.

The measure of team performance we used was the percentage of games won during the regular season. This is the same measure employed in all previous studies of succession in sports teams of any type. We rejected other indicators of team performance like numbers of assists, turnovers, or other more process-related variables for two reasons. First, as Scott (1977) indicated, outcome measures, if and when they are available and can be unambiguously defined, are better than process or structure measures because they get at the concept of effective performance more directly. One of the advantages of using sports teams in these analyses is that there is no problem of multiple constituencies wanting different outcomes from an organization. Sports teams have a clear measure of success—their win-loss records. Second, a focus on process requires investigators to develop a theory of how such processes are related to outcomes, and developing a theory of basketball is beyond the scope of this paper. Suffice it to note that the relationships between process and outcome measures are not straightforward. For instance, in our sample, the number of turnovers was correlated only .01 with the teams' win-loss percentages.

Following Brown (1982) who observed a large effect for the number of new players, we also included this variable in the analysis. New players were all players on a team's roster during a given season who had not been on the roster during the preceding season.

The ability of new coaches was measured in three ways. First, we measured cumulative win-loss records or percentages of games won prior to succession for all coaches who had previously coached in either the NBA or the ABA (American Basketball Association). This record included all partial as well as complete seasons coached; those with no previous experience received a value of zero for this variable. Of course, coaches are not totally responsible for their previous records as the ability of the players they have coached and the willingness of owners to spend money to attract good players are partly outside their control. Nevertheless, we presume an individual's past record to be a reasonable proxy for ability. Again, one of the advantages of studying sports teams is that the measures of ability are clearer and cleaner than in other settings.

Second, a dummy variable indicating whether coaches had previously coached professional basketball, either NBA or ABA, was used as a measure of ability gained through experience.

Third, in order to (1) account for the fact that coaches' win-loss records are not completely under their control and (2) to better isolate coaches' effects on team performance, we measured coaches' ability as the amount of improvement or decline in performance that occurred in teams coached prior to a succession event. Looking only at performance improvements or declines controls for relatively stable team characteristics such as owner resources and player ability and is therefore a somewhat clearer indicator of the ability of coaches to affect performance than previous win-loss records alone. For a given coach, performance improvements or declines were measured in two ways: (1) the percentage of seasons in which the win-loss record of the coach's team was better than that team's average win-loss record under its previous coach; and (2) the actual increase or decrease in win-loss record over the average win-loss record achieved by the previous coach. We averaged both measures over coaches' entire professional careers up to and including the seasons prior to successions. In a few cases, we could not compute these measures because coaches' first jobs had been with newly created teams or because their first jobs started before 1969 and lasted until 1977 or later and information about prior coaches was not available. In these cases, both measures received values of zero.

A dummy variable indicating succession was coded 1 if succession occurred either during the current season or between the prior and current seasons. To test the effects of coaches' records on succession, we created four interaction terms. The first multiplied the succession dummy variable by a coach's previous record; the second multiplied the succession dummy variable by the dummy variable assessing whether or not a coach had previous experience in professional basketball; the third multiplied the succession dummy by the percentage of seasons a coach produced performance improvements; and the fourth multiplied the succession dummy by the average change in win-loss record produced by a coach. If succession had not occurred, neither a coach's record nor experience was relevant or included in the analysis, since each would be multiplied by zero. The effect of prior experience or past record entered only in instances of succession.

RESULTS

Table 1 presents means, standard deviations, and a correlation matrix for all the variables. As the data combine time-series and cross-sectional observations, tests of significance for the correlations are difficult to interpret and are not reported (Hannan & Young, 1977). Table 1 shows that performance in the past season ($r = .52$) and the number of new players ($r = -.59$) have the strongest effects on performance in a current season. These results parallel those of Brown (1982) and of Allen and his colleagues (1979). Table 1 also reveals a positive effect for a coach's past record on performance ($r = .26$) and a small positive effect for a coach's having had experience in professional basketball ($r = .045$).

Table 1 indicates something about the conditions under which succession is most likely to occur. Teams that had poor records in the season just past are more likely to experience succession than teams that had good records ($r = -.305$), and they are also more likely to experience turnover on the roster ($r = -.45$). This replicates Brown's (1982) finding for professional football teams. The pattern of correlations also indicates why succession may appear to lead to a vicious cycle of performance decline. If succession and player turnover tend to occur together ($r = .26$), and player turnover leads to perfor-

TABLE 1
Means, Standard Deviations, and Correlations for All Variables

Variables	Means	s.d.	2	3	4	5^a	6^a	7	8
1. Percent of games won by team, current season	.505	.132	.524	.257	.045	−.059	−.006	−.591	−.105
2. Percent of games won by team, past season	.503	.122		.419	.100	−.244	−.110	−.451	−.305
3. Coach ability, past record	.436	.182			.844	.077	.097	−.239	−.551
4. Coach ability, previous professional coaching	.891	.313				.100	.052	−.080	−.614
5. Coach ability, percent of seasons improved over past coach^a	.068	.239					.732	.297	.501
6. Coach ability, change in win/loss record over past coach^a	.009	.053						.236	.262
7. Number of new players	6.727	3.013							.256
8. Succession between current and past season or within current season^b	.245	.432							

[a] Correlations with these variables should be interpreted with caution; see text for explanation.

[b] 1 = yes

mance decline (r = -.59), then observers may partly misattribute decrements in performance resulting from player turnovers to succession events. If decline in performance produces both further successions and further changes in a roster, performance may continue to decline even if succession itself has had no effect on it.

It is important to note that some of the relationships reported in Table 1 should be interpreted with extreme caution because the indicators of coach ability were not measured identically for teams that did not experience succession. In particular, we measured coaches' win-loss records and previous professional experience for all teams, whether or not they experienced succession. Therefore, these two measures of ability provide a relatively good indicator of the effect of ability on performance, regardless of whether or not succession occurs. However, because the measures of ability that track improvements in performance required locating data about coaches who were not in the sample of interest, we only calculated these measures for those teams in which succession actually occurred. Where succession did not occur, we coded these measures as zero. Therefore, because they are confounded with succession, these measures should not be used to look at the overall relationship between ability and performance.

TABLE 2
GLS Estimates of the Effect of Succession on Performance[a]

	Equations				
Variables	**1**	**2**	**3**	**4**	**5**
Performance in past season	.237****	.318****	.363****	.241***	.310****
	(2.57)	(5.31)	(8.47)	(2.68)	(10.38)
Number of new players	−.020****	−.020****	−.020****	−.021****	−.021****
	(5.72)	(16.64)	(14.36)	(5.92)	(38.59)
Succession	.038	—	—	—	—
	(1.69)				
Succession × coach's past record	—	.093****	—	—	—
		(3.88)			
Succession × previous professional coaching	—	—	.050****	—	—
			(5.64)		
Succession × percent of seasons improved over past coach	—	—	—	.087**	—
				(2.16)	
Succession × change in win/loss record over past coach	—	—	—	—	.546 ****
					(4.81)
Intercept	.509	.460	.428	.516	.481
R^2	.436	.444	.447	.451	.437

[a] Dependent variable is performance in the current season. Unstandardized regression coefficients are reported and t-values appear in parentheses.

*p<.10
**p<.05
***p<.01
****p<.001

It is also important to remember that measuring ability only for succession events has no effect on the subsequent analyses that we report. As noted earlier in this paper, we multiplied ability by the succession dummy variable, which was not zero only when succession had actually occurred.

Table 2 presents the results of five multiple regression equations.[2] The first equation included only the analysis for effects of succession, without reference to either experience or the past records of new coaches. The equation indicates that both past performance and player turnover have significant effects on current performance, but there is no significant effect for succession. The absence of a succession effect replicates the predominant finding in the literature.

---

[2] An alternative method for estimating the equations is to use ordinary least squares and to include dummy variables for each team (least squares with dummy variables). When this approach was used for the analysis reported here, the results were essentially the same as those reported in Table 2 with the following three exceptions. First, performance in the past season had no effect on performance in the current season in any of the equations. This result is expected because the effects of the lagged dependent variable are essentially captured by the dummy variables for each team. The dummy variables for Chicago, Cleveland, Detroit, San Diego, and Utah were significant and negative in most of the equations. The variables for Los Angeles and Philadelphia were significant and positive in all of the equations. All of the other dummy variables were not significant in any of the equations. Washington was the omitted category. Second, the effects of the four interaction terms are somewhat weaker, although all are still positive and most are still significant. Third, R^2 increases by about .10 in all of the equations.

The next four columns indicate that the quality of new coaches has a significant effect on performance. Equation 2 estimates the effect of succession on performance with the past record of the coaches considered; equation 3 takes into account the coaches' previous professional experience; and equations 4 and 5 consider improvements or declines in past performance brought about by coaches. In all instances, the interaction term is statistically significant and in the expected direction, indicating that (1) coaches with better prior win-loss records, (2) coaches with previous experience, and (3) coaches who have improved the performance of other teams are all associated with better team performance after succession. Because only those coaches with previous experience have records to be entered into the equation, there is too much collinearity between the various interaction terms to make it sensible to enter more than one at a time. Thus, although we cannot estimate which of the effects is predominant, we can conclude that the ability and experience of a coach have an effect on a team's performance after succession.

DISCUSSION

It seems increasingly clear that succession is a complex phenomenon with multiple effects depending on timing, the choice of inside or outside successors, and the abilities of successors. Previous research has indicated, for instance, that outside succession is likely to be associated with more organizational change than inside succession (Carlson, 1962; Helmich & Brown, 1972). In their study of baseball teams, Allen and his colleagues (1979) found some evidence for midseason succession being more disruptive than end-of-season succession and for inside succession being less disruptive than outside succession. Results of the present study indicate that, as common sense might suggest, the ability of a successor affects a team's performance after the succession.

The findings of both previous and the present research suggest that merely knowing that succession has occurred is not an adequate basis for predicting its consequences. Rather, the type of succession, when it occurs, and particularly, the abilities of replacement managers all need to be accounted for in assessing likely effects of succession. In other words, it is probably not fruitful to try to find a single succession effect because the consequences of successions are likely to vary dramatically depending on the conditions surrounding them. This research replicated the finding that merely knowing that succession has occurred does not increase ability to predict subsequent organizational functioning. Results also confirmed the hypothesis that the ability of a new manager is in fact an important predictor of organizational performance after succession. Thus, consistent with previous findings (Smith et al., 1984), this research suggests that it is important for organizations to select individuals with the best performance records from pools of available replacements.

This research also helps to disentangle the two theoretical problems mentioned in the beginning of this article that are associated with succession. Some past researchers who have studied the effects of succession events themselves have suggested that finding no effect of succession indicates that leaders cannot significantly affect organizational performance. This research indicates that the question of whether leaders affect organizational performance cannot be answered by merely looking at organizational performance before and after succession. Instead, investigators should examine conditions surrounding successions and the characteristics of successors, particularly ability and experience. By examining these aspects of succession, future research may be able to identify types of succession situations that allow leaders to have substantial influence on organizational outcomes and types of succession situations in which leader influence is more limited. Similarly, future research may be able to identify characteristics in addition to ability and experience that affect leaders' abilities to influence organizational outcomes.

REFERENCES

Allen, M. P., Panian, S. K., & Lotz, R. E. 1979. Managerial succession and organizational performance: A recalcitrant problem revisited. ***Administrative Science Quarterly***, 24: 167–180.

Allen, M. P., & Panian, S. K. 1982. Power, performance, and succession in the large corporation. ***Administrative Science Quarterly***, 27: 538–547.

Brown, M. C. 1982. Administrative succession and organizational performance: The succession effect. ***Administrative Science Quarterly***, 27: 1–16.

Carlson, R. O. 1962. ***Executive succession and organizational change***. Danville, Ill.: Interstate Printers and Publishers.

Eitzen, D. S., & Yetman, N. R. 1972. Managerial change, longevity, and organizational effectiveness. ***Administrative Science Quarterly***, 17: 110–116.

Gamson, W. A., & Scotch, N. R. 1964. Scapegoating in baseball. ***American Journal of Sociology***, 70: 69–76.

Gephart, R. P., Jr. 1978. Status degradation and organizational succession: An ethnomethodological approach. ***Administrative Science Quarterly***, 23: 553–581.

Gouldner, A. W. 1954. ***Patterns of industrial bureaucracy***. Glencoe, Ill.: Free Press.

Grusky, O. 1963. Managerial succession and organizational effectiveness. ***American Journal of Sociology***, 69: 21–31.

Grusky, O. 1964. Reply to scapegoating in baseball. ***American Journal of Sociology***, 70: 72–76.

Guest, R. H. 1962. Managerial sucession in complex organizations. ***American Journal of Sociology***, 68: 47–54.

Hannan, M. T., & Young, A. A. 1977. Estimation in panel models: Results on pooling cross-sections and time series. In D. R. Heise (Ed.), ***Sociological methodogy***: 52–83. San Francisco: Jossey-Bass.

Helmich, D. L., & Brown, W. B. 1972. Successor type and organizational change in the corporate enterprise. ***Administrative Science Quarterly***, 17: 371–381.

Lenz, R. T. 1981. "Determinants" of organizational performance: An interdisciplinary review. ***Strategic Management Journal***, 2: 131–154.

Lieberson, S., & O'Connor, J. F. 1972. Leadership and organizational performance: A study of large corporations. ***American Sociological Review***, 37: 117–130.

March, J. C., & March, J. G. 1977. Almost random careers: The Wisconsin school superintendency, 1940–1972. ***Administrative Science Quarterly***, 22: 377–409.

Pfeffer, J. 1977. The ambiguity of leadership. ***Academy of Management Review***, 2: 104–112.

Salancik, G. R., & Pfeffer, J. 1977. Constraints on administrative discretion: The limited influence of mayors on city budgets. ***Urban Affairs Quarterly***, 12: 475–498.

Salancik, G. R., & Pfeffer, J. 1980. Effects of ownership and performance on executive tenure in U. S. corporations. ***Academy of Management Journal***, 23: 653–664.

Scott, W. R. 1977. Effectiveness of organizational effectiveness studies. In P. S. Goodman, J. M. Pennings, & associates (Eds.), ***New perspectives on organizational effectiveness***: 63–95. San Francisco: Jossey-Bass.

Smith, J. E., Carson, K. P., & Alexander, R. A. 1984. Leadership: It can make a difference. ***Academy of Management Journal***, 27: 765–776.

The Sporting News. 1978–82. ***The Sporting News Official NBA Guide***. St. Louis, Mo.

Weiner, N., & Mahoney, T. A. 1981. A model of corporate performance as a function of environmental, organizational, and leadership influences. ***Academy of Management Journal***, 24: 453–470.

Jeffrey Pfeffer received his Ph.D. degree from Stanford University. He is currently Professor of Organizational Behavior at the Graduate School of Business, Stanford University. His research interests include organizational demography, wage structures in organizations, and the structuring of the employment relationship.

Alison Davis-Blake received a Master of Organizational Behavior degree from Brigham Young University and currently is a doctoral candidate at the Graduate School of Business, Stanford University. Her research interests include organizational labor markets, the determinants of organizational salary structures, and the interorganizational diffusion of personnel practices.

© *Academy of Management Journal*
1986, Vol. 29, No. 1, 84–100.

EFFECTS OF ALTERNATIVE JOB EVALUATION METHODS ON DECISIONS INVOLVING PAY EQUITY

ROBERT M. MADIGAN
Virginia Polytechnic Institute and State University
DAVID J. HOOVER
Miami University, Ohio

This study compares results obtained from six different job evaluation methods in terms of (1) their effects on job classification decisions and (2) inferences regarding the equity of job hierarchies. Evaluations of 206 state-classified jobs indicated that decisions about job hierarchies and estimates of the extent of gender-based discrimination in classifying jobs are dependent on the methods used.

The issue of pay equity for women has become a major social and economic issue in our society. A wide gap between the median earnings of men and women has persisted in the face of profound changes in women's participation in the labor force (Norwood, 1982). This has fueled a national debate over pay discrimination that was heightened considerably by the conclusions of a National Academy of Sciences (NAS) panel that various aspects of wage determination, particularly job evaluation procedures, are susceptible to gender based discrimination (Treiman & Hartman, 1981). The NAS report also suggested that modified job evaluation procedures could provide standards of relative job worth that might uncover or reduce discrimination. Such standards would go beyond the historical rule of equal pay for equal work by mandating equal pay for dissimilar jobs of comparable worth as measured by internal or nonmarket criteria of job value. Advocates of the comparable worth principle of pay determination argue that it offers a means of significantly and immediately reducing the pay gap between men and women. If this principle were public policy, many more situations in which there is pay disparity would become illegal.

Although arguments against adopting the comparable worth principle are formidable (Livernash, 1980; Lorber, Kirk, Samuels, & Spellman, 1985; Nelson, 1980), support for it is burgeoning. State and local governments have launched more than a 100 different comparable worth initiatives (Mayer, 1984), including a 1984 Minnesota statute mandating implementation of the comparable worth principle in each political subdivision in the state. Interestingly, few of the initiatives specify the criteria of job worth or the measurement procedures to be used in establishing relative job worth. This

---

The authors would like to thank H. Patricia Curran from the Michigan Department of Labor, Office of Women and Work for her support; thanks also to the anonymous reviewers for their many helpful comments and suggestions.

omission has potentially significant consequences. Because universal norms of worth do not exist (Mahoney, 1983), the concept of worth incorporated in different job evaluation plans is likely to vary. In addition, the psychometric qualities of evaluation plans vary (Madigan, 1985). Consequently, relative job worth is likely to be a function of the method of job evaluation used; this likelihood has important legal and financial implications as well as implications for employee-management relations.

If decisions on pay classification are a function of the method used to evaluate jobs, it follows that inferences regarding pay discrimination also depend on method. In analyses of pay discrimination, job evaluation procedures define variables for job worth or job level. The studies of pay equity or comparable worth that most states have conducted or considered are particularly susceptible to method differences because job worth is their sole criterion of the fairness of pay systems. A recent court decision, American Nurses Association v. Illinois (1985), explicitly recognized the possibility that different job evaluation methods could produce different results. This decision disallowed a job evaluation study as the basis for a pay discrimination action. On the other hand, in the much publicized case of AFSCME v. State of Washington (1983), a simple regression analysis in which pay rates and evaluation points were entered was critical to the union's case. In effect, the court rejected the state of Washington's definition of worth in favor of a consultant's job evaluation study. It is interesting to speculate how the Washington court would have ruled had advocates submitted competing measures of job worth for consideration.

The research literature on correspondence among job evaluation plans is sparse and subject to differing interpretations. Snelgar (1983) reviewed five studies and concluded that different methods produce highly similar results. Schwab (1980) emphasized divergence in results on the basis of his review of four studies, two of which Snelgar had cited. It is noteworthy that none of the reported studies assessed the extent of correspondence between methods in terms of decisions about job classifications. Rather, correlation coefficients provided their indices of agreement, an approach that can mask considerable divergence in actual classification decisions that are based on different methods of evaluating jobs.

Hence, the purpose of this study was to investigate two related questions: (1) To what extent do pay structures vary by job evaluation method? (2) Are the results of studies of pay equity or comparable worth sensitive to differences in job evaluation methods?

METHODS

The data analyzed here were collected for an earlier study of pay equity in the state of Michigan civil service job classification structures; the senior author served as a member of the technical advisory committee for that study. The primary purposes of the original study were (1) to assess the feasibility of developing and using an approach that would evaluate all types of jobs within the state of Michigan on a common scale of value; and (2) to analyze the extent to which jobs dominated by members of one gender received equal pay for equal value as measured by job evaluation. In recognition of the potentially discriminatory aspects of conventional job evaluation procedures (Treiman, 1979), project specifications required controls for possible gender effects in both job analyses and in the development and application of the job evaluation plan.

The earlier study was conducted during 1980 and 1981 by a management consulting firm under the direction of the Office of Women and Work, Michigan Department of Labor. The data collection procedures described below summarize relevant portions of the project report. This study appears to have been unique among the many recent public sector analyses of pay equity in that investigators applied two distinct methods of job evaluation to a job sample. However, the study did not directly address the issue of differences among job evaluation methods. Hence, for the present study we developed alternative procedures for scoring job worth and analyzed data from the original study again.

Job Sample

Jobs selected for the original study were drawn from 206 job classifications in seven service groups in the state's classification system. The groups were engineering and science; law enforcement; labor and trades; clerical support; public safety, security, and regulatory work; domestic work; and human service. The study excluded the legal, executive/administrative, and physician groups. Job classes were selected to ensure that the sample (1) represented the seven service groups and their full ranges of job complexity, and (2) contained jobs that represented groups' gender composition fully enough to permit statistical analysis. Classes of jobs that had between 30 percent and 70 percent men or women were designated as mixed classes; those with over 70 percent of one gender were defined as dominated by men or women. The 206 classes in the study took in approximately 25 percent of the total active job classes in the seven service groups. The percent of job classes sampled within groups ranged from 22 to 48 percent. The imbalance in gender composition among service groups necessitated these intergroup sampling differences. Over half the job classes consisted of jobs held mainly by men. In addition, most job classes were dominated by one or the other gender. In the final sample, men dominated 69 classes, women dominated 84, and the remaining classes were mixed.

Job Analysis

Eight job analysts were hired for the original state of Michigan project. All analysts were currently enrolled in graduate programs, or had been enrolled recently, and all had completed at least master's degrees in industrial/organizational psychology or industrial relations. Analyst training consisted of a two-day program covering interview procedures, use of a patterned interview guide, job evaluation procedures, and project documentation requirements. During the job evaluation phase of the training, analysts specialized in using one of the two rating instruments, the Position Analysis Questionnaire (PAQ) (McCormick, 1979) or a traditional point plan, both of which are discussed later in this paper. Analyst teams with one man and one woman were then formed to control for possible gender-related variance in perceptions of job characteristics (Arvey, Passino, & Lounsbury, 1977).

Job analysis procedures were designed to ensure that job information would be identical for both evaluation methods. Analysts collected job information by reviewing descriptions of both generic classifications and individual positions. Job incumbents were then interviewed. One interview was conducted for each of the 206 classes. A typical interview included four analysts—one PAQ and one traditional team—and from two to eight jobholders with at least six months tenure in the same class. The number of people in a class and the apparent diversity of assignments within it determined the number of interviewees per class. Analysts tape-recorded all sessions to provide complete records of the interviews for later reference and to serve as primary input when one team could not be present.

Interviews followed a semistructured procedure. An interview guide developed by the consultants specified the content and sequence of information collection to ensure that interviewers would base job ratings on comprehensive, job-related information.[1] The guide specified certain key questions critical to accurate completion of the evaluation instruments. Within these constraints, interviewers developed their own questions and lines of inquiry. They took notes during interviews and were instructed to carefully document their observations at the conclusion of each session.

Data collection procedures differed from typical job evaluation practices in that analysts did not observe actual work situations. Moreover, they were less knowledgeable than internal analysts would be about all the types of work in the service groups. Thus, the job information collected could have been deficient or misevaluated to some degree. These shortcomings would be serious for an analysis of equity in pay structures if the objective were to establish a job hierarchy that accurately reflects differences in job content. However, the present study did not address the question of equity in the Michigan classification system. Rather, its focus was correspondence between evaluation methods. The critical consideration here was whether job information was the same for all methods. The original study's data collection procedures were appropriate to this objective.

Job Ratings

Analysts completed evaluation instruments as soon as possible after conducting interviews. Analysts were instructed to refer to their tape recordings and notes prior to completing ratings if delays occurred. Each analyst was asked to work independently in order to facilitate reliability checks. They submitted the four completed job evaluation instruments, interview notes, and tape recordings to the project staff when an evaluation was completed.

---

[1] The interview guide and other instruments are available from the project archives maintained by the Office of Women and Work, Michigan Department of Labor.

Analysts' efforts underwent a number of quality checks. Project consultants closely monitored the early stages of data collection and completed third sets of ratings on 50 jobs to provide an additional means of auditing analysts' accuracy and reliablity. They also reviewed an early sample of each analyst's job evaluations for evidence of range restriction and leniency or severity.[2] When job evaluations were completed, project staff checked the instruments submitted by each pair of analysts for discrepancies in item scores of more than one level. Items with such descrepancies went back to the analysts for comparison with notes and tape recordings. As a final step toward high measurement quality, the two ratings on each traditional factor and PAQ item were averaged to provide a single set of ratings for each job on each instrument. Hence, the two basic sets of job ratings used in the study consisted of (1) the mean PAQ item scores, and (2) the mean, unweighted, traditional-plan factor-level scores.

Instruments

Traditional plan. The traditional plan was a modified version of the primary standards of the federal government Factor Evaluation System (FES), a method designed to apply to a broad range of nonsupervisory positions. The FES method rates each job on a set of nine job worth or compensable factors of the type common to traditional point-factor plans. These include knowledge required by a job, its physical demands, complexity, work environment, and so forth. Compensable factors receive point values that reflect their relative importance to a job's worth or value as defined by an organization. The sum of the ratings on each factor constitutes total job worth, and jobs are ordered in the pay structure according to these total scores.[3]

The FES plan primary standards were modified in three ways. First, the project team added factors to evaluate supervisory responsibilities and characteristics related to job stress. Second, FES scale definitions and examples were revised to make them more relevant to the job sample. Finally, a new factor-weighting scheme was developed. These changes reflected the judgment of an advisory panel consisting of representatives of unions, civil service, the Equal Employment Opportunity Commission, and women's rights groups. In the opinion of this broadly representative group, the final set of 11 factors was an appropriate and nonbiased set of job worth criteria. This process of developing a job evaluation method followed suggestions of the NAS and other critics of contemporary job evaluation plans (Remick, 1984).

Position Analysis Questionnaire. The PAQ is a 187-item quantitative job analysis instrument designed to encompass every human behavior that might arise on the job (McCormick, 1979). Item scores are combined to produce dimension scores (McCormick, Jeanneret, & Mecham, 1977) that can be treated as compensable factors in job evaluation applications. Dimension scores are converted to measures of job worth by using multiple regression to find the combination of weighted dimension scores that most accurately captures a criterion structure, which is typically the wage rates for benchmark jobs. The resulting least-squares line becomes a policy line for setting wage rates, and observed dimension scores for all jobs are then entered into the equation to obtain job values. This policy-capturing approach to job evaluation is not unique to the PAQ. Historically, most evaluation plans have mirrored predetermined job hierarchies, and statistical weighting has been used extensively for over 35 years.[4] The PAQ differs from past approaches to job evaluation primarily in the number and specificity of its questionnaire items and in its direct application of a regression equation to evaluate job worth. Obviously, to the degree a criterion structure reflects bias, this evaluation procedure captures such bias. In the present study, two different criterion structures were used to develop PAQ-based equations for job worth. These are discussed in the next section.

---

[2] Results of these early reviews for discrepancies and problems were not available.

[3] See Henderson (1985) or Treiman (1979) for a more complete discussion of the FES plan and construction of point plans.

Scoring Procedures

The study used six methods of developing relative worth scores from the two basic sets of job ratings. Methods 1 and 2 used subjective weights; methods 3 and 4 were statistically weighted versions of traditional plan ratings; and methods 5 and 6 were the two PAQ equations. The six scoring procedures were as follows:

Method 1. The traditional plan factors and scale levels received point values in accordance with both the factor weights determined by the advisory panel and with normal procedures for point plan development (Henderson, 1985). Because Factors 8 through 11 consisted of two subfactors, this procedure yielded a total of 15 factor scores for each job. We will refer to these factor scores as scale scores. The sum of scale scores was determined for each of the 206 jobs to establish its total job worth and relative position in the job hierarchy. Table 1 presents the factor weights for methods 1, 3, and 4. Method 2 weights were the same as those of method 1 for the selected factors.

Method 2. A simplified version of method 1 was developed in which only four of the original scales were used. Research results have often indicated that two to four factors dominate point evaluation plans (Lawshe, 1945; Lawshe, Dudek, & Wilson, 1948; Madigan, 1985). The traditional plan was no exception. Factor analysis revealed a three-factor solution explaining 93 percent of common variance in unweighted scale level ratings, or a four-factor solution explaining 99 percent. We developed a four-factor simplified plan with four scales from the traditional plan selected as proxies for the underlying factors primarily on the basis of their factor loadings (Table 2).

---

[4] See, for example, Stieber's (1959) discussion of the origins and development of the steel industry's job evaluation plan.

TABLE 1
Factor Weight Allocations for Evaluation Methods 1, 3, and 4

Traditional Plan Factors	Method 1	Method 3	Method 4
Knowledge required	11.2%	30%	7%
Supervisory controls	9.8	1	5
Performance guidelines	9.2	5	3
Complexity	10.0	2	11
Output effect	10.9	1	9
Personal contacts	7.0	5	11
Purpose of contacts	8.7	1	1
Physical demands			
Abilities	4.0	1	1
Exertion	4.4	1	17
Work environment			
Risk	4.5	3	7
Working conditions	3.9	2	11
Stress conditions			
Workpace	3.5	5	5
Interruptions	4.0	3	1
Supervisory responsibilities			
Degree of responsibilities	5.4	25	10
Number supervised	3.5	15	1
	100.0%	100%	100%

TABLE 2
Oblique Factor Pattern Matrix for Traditional Plan Scale Scores

Scale Scores	Factor 1	Factor 2	Factor 3	Factor 4
Knowledge required	**.86**	.06	.00	.04
Supervisory controls	**.71**	.09	.26	.04
Performance guidelines	**.74**	.06	.09	.00
Complexity	**.80**	.04	.10	.07
Output effect	**.76**	.09	−.06	.09
Personal contacts	**.78**	−.14	−.09	.14
Purpose of contacts	**.84**	.12	−.05	−.01
Physical abilities	−.24	**.42**	−.10	.17
Physical exertion	−.01	**.76**	.00	−.02
Work risk	.12	**.89**	.05	.01
Working conditions	.10	**.74**	.04	−.11
Workpace	.19	.00	.03	**.66**
Interruptions	.10	.06	.14	**.75**
Supervisory responsibilities	.03	.02	**.96**	.01
Number supervised	.06	.01	**.95**	.08
Eigenvalue[a]	6.25	2.18	1.29	.71
Percent variance	41.70	14.50	8.60	4.70

[a] Initial factors.

More elegant approaches to developing factor scores could have been used; however, the objective here was not to optimally measure an underlying commonality but to create a simple plan similar to those that are in use. Ratings from the first seven scales all loaded heavily on the first common factor. The first scale, knowledge requirements, was selected to represent this factor. In a similar fashion, the scale for working conditions was chosen for Factor 2, supervisory responsibilities represented Factor 3, and work place interruptions represented Factor 4.

Method 3. A statistically weighted or policy-capturing version of the traditional plan was developed with the maximum of the pay range for mixed job classes as the criterion structure. One of the criticisms of factor weighting procedures has been that the benchmark jobs used to determine weights are typically jobs dominated by men, which relegates factors associated with jobs held predominantly by women to minor importance (Treiman, 1979). Method 3 is one possible response to this criticism; the rates assigned to the mixed classifications are assumed to represent a fair or at least gender-blind pay policy that the job evaluation model captures. This assumption is acceptable here, but an actual equity analysis would need to establish that a given criterion structure represented the entire domain of work in a job structure.

Researchers for the original study chose to use maximum pay ranges rather than midpoints because many employees were at or near the maximums. Thus, the analysis of equity in the classification structure also served as a rough indicator of equity in actual pay distribution.

Factor weights in method 3 were based on beta weights from a multiple regression equation in which maximum pay and unweighted traditional plan scale scores were entered.[5] These weights differed from the subjectively determined weights used in method 1. In method 1 fair weights were subjectively determined via a consensus process, whereas in method 3, job ratings were used to statistically model a fair criterion structure and factor weights came from the beta weights that resulted. In both cases, the point values allocated to factors and scale levels reflected the weighting schemes.

---

[5] Statistically assisted weighting of point factor plans is an indirect type of policy-capturing procedure. Beta weights are used to help allocate values to the factors and degree levels, rather than for weighting of observed scores as in the PAQ procedure. In point plans, the maximum point values assigned to each factor reflect the relative importance of factors (Sargent, 1972). The weighting task involves determining a point allocation scheme that generates a point hierarchy closely corresponding to the criterion structure when it is used to evaluate the benchmark jobs. In small organizations, this is typically a trial-and-error process. When a large number of jobs is involved, beta weights can facilitate the process because they indicate factors' relative contributions to variance in the criterion structure. However, beta weights are not readily transformed into an ideal weighting scheme, particularly if a high degree of multicollinearity exists among compensable factors. The procedure used here was to proportionally reduce the beta weights to create a 100-point distribution, retaining a minimum weight of one for factors. This process established the relative importance of factors. The final set of numerical weights was determined experimentally by testing variations of the basic weighting scheme for their affect on R^2. The usual test for a weighting scheme is its hit rate (Gomez-Mejia, Page, & Tornow, 1982), but the multiple classification structures in this job sample precluded this approach.

Method 4. This scoring procedure differs from method 3 only in the criterion structure used to derive factor weights. Pay rates of classes dominated by men were used as the criterion structure to test for the problem noted in the description of method 3. Thus, this method determined the extent to which a different and possibly biased set of benchmark jobs affects job evaluation results.

Method 5. As noted above, two different PAQ equations were developed. In method 5, this study used overall PAQ dimensions to capture maximum pay for the the total job sample. Because classification rates were heavily based on pay survey data, this criterion structure reflects market value. To the degree jobs held by women were undervalued in the marketplace, this equation was not gender-blind. However, the large and disparate job sample provides a possible rationale for defining this as a fair criterion. More important for the present study, the large sample equation roughly approximates the approach advocated by PAQ proponents for use of this instrument in job evaluation applications (McCormick, Mecham, & Jeanneret, 1977; Mecham, 1983). Overall dimension scores were selected because (1) they are a more concise and less redundant summary of PAQ information than the division dimensions; (2) the two sets of scores were equivalent in terms of the proportions of variance in pay rates explained; and (3) overall interrater dimension reliability estimates were slightly better. The mean reliability coefficient for the overall dimensions was .83, slightly higher than that of the traditional plan factor scores (.80).[6]

Method 6. This scoring procedure differed from method 5 only in that classes dominated by men provided the criterion structure used to derive dimension coefficients. Results from this method allowed for comparisons between methods with the same benchmark structures but different factor contents and treatments of observed scores.

In summary, the job evaluation scoring methods (1) were based on two different job rating procedures, a traditional plan and the PAQ; (2) represent two basic approaches to compensable factor weighting, judgmental and statistical; and (3) reflect the effects of using different criterion structures in the statistical weighting procedure.

Analysis

Convergence among methods of job evaluation was analyzed in two ways. First, we calculated intermethod correlation coefficients to provide a basis for comparisons with earlier studies. Second, job worth scores from methods 2 through 6 were transformed to the metric of method 1 (Nunnally, 1967:107) to allow direct comparisons between job scores of different methods. Method 1 had a 500-point theoretical maximum with actual scores ranging from 131 to 416, a range of 285 points. We calculated differences in job scores between pairs of methods for all 206 classes and all possible method pairs. Means and standard deviations of the distributions of absolute difference scores were then calculated to provide indicators of disagreement between evaluation methods. Meaningful interpretation of point differences, of course, requires knowledge of the point ranges associated with pay grades. For example, in the Midwest Industrial Management Association plans, pay grade increments equate to point differences of 6 to 8 percent of the total range of scores. That is, a point difference between jobs of greater than 8 percent will result in different pay grades. In the present study, we used a 10 percent criterion to estimate agreement rates; a point differential between methods of less than 10 percent of the total range of job scores constituted agreement. This 10 percent or ten-level rule provided a liberal definition of classification agreement; it was apt to overstate the agreement rates because pay structures for service groups each contained from 8 to 13 pay grades.

---

[6] These reliabilities compare favorably with those of previous job evaluation studies.

The effects of method differences on analysis of equity in a pay structure were determined by using the alternative job worth measures in two basic approaches to discrimination analysis suggested by the NAS report (Treiman & Hartman, 1981). First, we estimated the effect of the gender composition of job classes on the classification rates by regressing maximum pay on job worth points and a gender variable. A negative, significant gender coefficient after job worth is controlled for indicates potential discrimination. Separate analyses were conducted for each job evaluation method using both categorical and continuous measures of gender composition; in the first, the gender variable was coded woman-dominated = 1; in the second, the variable was the ratio of women to total incumbents.

Second, the pay rates for job classes dominated by women were compared with estimates of fair pay for these classes. We developed estimated or predicted fair pay levels by (1) regressing maximum pay on scale/dimension scores of job classes dominated by men, and (2) applying the resulting coefficients to the observed scores of jobs held mainly by women. The rationale for this analysis is that fairness dictates that compensable factors for all job classes should be equally valued. Hence, when estimated rates for women exceed actual rates on the average, there is evidence of possible discrimination. Again, we conducted six separate analyses, testing mean differences between actual and predicted pay rates for significance with paired-comparison t-tests.

Previous research has documented the potential problems with using regression procedures to detect or correct pay discrimination (Ramsey, 1979; Roberts, 1980). However, such issues are not germane to the present study, which focused on the effects of evaluation methods, not on analysis of classification discrimination. The procedures described in the previous sections were appropriate because they represented current practices in analyzing pay equity, and the issue under study was the degree to which the outcomes of such analyses vary by evaluation method.

RESULTS

Table 3 presents intercorrelations among job evaluation scores generated by the six methods. This table also shows the simple correlations between job scores and maximum pay for the 206 classifications as well as the multiple correlations between maximum pay and sets of factor and dimension scores. Results indicate considerable disparity in outcomes of job evaluation plans. The low correlation between methods 4 and 6 (.41) is particularly notable because (1) the same criterion structure was the source of factor weights in both plans, and (2) the two methods explained the same percentage of criterion variance. The extent of disagreement between these two plans reveals the potential effects of evaluating jobs with different sets of job worth criteria and scoring procedures. It is also worth noting that multiple correlations between maximum pay and PAQ dimension scores are no higher than correlations between maximum pay and the traditional plans' scores. In this job sample the detailed level of analysis required by the PAQ yielded no advantage in explaining variance in classifications.

Table 4 provides a basis for evaluating differences in methods in terms of their potential effects on classification decisions. The table shows mean differences in job worth scores and estimated agreement rates for each pair of methods. Differences between methods were substantial. For example, the two basic procedures (1 and 5) differed markedly, with probable agreement on only 54 percent of the jobs. The differential in job worth scores between methods 1 and 5 was less than 10 percent of the range of job scores for 54 percent of the jobs.[7] If the agreement criterion had been relaxed to 15 percent, divergence in classification decisions would still have been 28 percent. Moreover, the size of many of the standard deviation values suggests that classification differences often would be greater than one level.

TABLE 3
Intercorrelations among Alternative Job Evaluation Scores and Classifications' Pay Rates[a]

Job Evaluation Methods	1	2	3	4	5	6
1. Traditional		.92	.90	.93	.83	.64
2. Simplified			.95	.86	.83	.63
3. Traditional — mixed				.84	.84	.63
4. Traditional — man-dominated					.71	.41
5. PAQ — total						.82
6. PAQ — man-dominated						
Maximum pay	.80	.81	.81	.68	.86	.71
R[b]	.89	.86	.89	.89	.86	.78

[a] n = 206 job classes.
[b] Multiple correlation, maximum pay for classification with set of factor/dimension scores.

---

[7] The range of job scores was defined as the mean range of the two methods compared.

The agreement rates in Table 4 might be compared with the findings of a recent comparison of the predictive accuracy of seven job evaluation methods (Gomez-Mejia, Page, & Tornow, 1982). That study operationally defined accuracy in terms of hit rates, the percentage of classifications made with a job evaluation method that were within ±1 grade of a job's actual assigned grade. The rates ranged from 49 to 73 percent, slightly higher than the current study's. However, two differences between the studies must be considered in drawing comparisons. First, Gomez-Mejia and colleagues looked at only one occupational family, management, and based all seven of their scoring procedures on a single set of incumbent responses to a customized 235-item questionnaire. These conditions could be expected to increase correspondence between methods. Second, agreement rates and hit rates are not directly comparable. The agreement rates in Table 4 merely indicate the extent to which two methods produced similar scores, not whether results fell into predetermined ranges of values.

Method differences also can significantly affect conclusions regarding the equitability of a pay structure. Table 5 displays results obtained when we used job scores from the six methods in the two approaches to analysis of discrimination in classification structures that the NAS report (Treiman & Hartman, 1981) suggested. The regression coefficients[8] indicate gender effects

TABLE 4
Differences in Job Scores and Classification Agreement between Pairs of Job Evaluation Methods[a]

	Job Worth Point Difference[b]		
Evaluation Method Pairs	Means	s.d.	Classification Agreement Rates[c]
1 v. 2	19.7	14.4	73%
1 v. 3	21.5	17.0	68
1 v. 4	18.0	13.5	78
1 v. 5	28.1	21.6	54
1 v. 6	41.0	31.2	40
2 v. 3	14.2	11.8	84
2 v. 4	26.6	17.8	55
2 v. 5	27.5	20.7	53
2 v. 6	42.0	30.8	34
3 v. 4	27.6	20.3	58
3 v. 5	27.1	21.3	55
3 v. 6	43.3	28.9	32
4 v. 5	36.8	28.0	44
4 v. 6	53.9	38.2	27
5 v. 6	28.2	23.1	54

[a] n = 206 job classes.
[b] Absolute difference.
[c] Agreement defined as a differential within 10 percent of the range of job scores.

---

[8] Regression coefficients were obtained using the dummy and ratio measures described in the previous main section.

after total job worth, as measured by the job evaluation methods, was controlled for. In all six analyses the coefficients were negative and statistically significant. However, the issue of interest here is the variance in discrimination indices across plans.

The regression coefficients for method 6 represent an improbable case—use of an auditing tool for analysis of discrimination derived directly from a criterion structure limited to jobs held mainly by men. Hence, method 6 should be disregarded here, except as evidence of what could happen if compensable factors do not represent a total job structure. However, substantial differences remain among the other five methods. For example, the monthly salary decrement associated with a 1 percent increase in the number of women in a job class is 64 percent greater for method 3 than for method 4 (1.92 vs. 1.17). The mean-differences tests show similar method effects. Analysis using method 1 suggests that, on the average, women-dominated classifications are underpaid by $427; the same analysis using method 5 suggests only a $57 discrepancy. In short, analyses of equity in pay structures are highly sensitive to differences in job evaluation methods.

DISCUSSION

The results reported here indicate that concern about the effect of differences in job evaluation methods on decisions pertaining to pay equity is well placed. If relative job worth is based solely on job evaluation procedures, choice of evaluation method will influence pay structures. Consequently, inferences regarding gender-based discrimination drawn from studies of pay equity based on job evaluation procedures are also method-dependent.

The results of this study have significant implications for voluntary and mandated attempts to incorporate the comparable worth principle in pay policies. As noted earlier, comparable worth initiatives tend to ignore the fact that there is no definitive test of the fairness of a pay structure. Equity in

TABLE 5
Comparison of Results from Discrimination Analyses Using Different Job Evaluation Scores[a]

	Job Evaluation Methods					
Discrimination Indicator	**1**	**2**	**3**	**4**	**5**	**6**
Regression coefficients[c]						
Percent women	−1.57	−1.81	−1.92	−1.17	−1.61	−4.68
Dummy variable[b]	−109.00	−138.00	−151.00	−93.00	−104.00	−309.00
Mean differences test						
Estimated v. actual	$427	$320	$137	$303	$57	$322
t[c]	23.3	17.5	7.1	15.8	3.0	14.3

[a] n = 206 job classes.

[b] Women = 1.

[c] All significant at the .01 level.

a pay structure is like beauty—in the eye of the beholder. The NAS report (Treiman & Hartman, 1981:96) explicitly recognized this fact in recommendations for removing or controlling sources of gender-related bias in the development and administration of evaluation plans. However, the NAS panel avoided the basic issue: in the absence of a complete consensus, whose values should prevail? Development of job evaluation procedures involves discovering and reconciling competing criteria of job value; individual values or biases are integral to judgments about the worth of jobs. Modifying the development and administration of evaluation plans might reduce bias but would not ensure that one gender-blind method is more equitable than another. The present study suggests that organizations could generate ranges of equitable pay structures and choose methods that best meet their objectives. Similarly, courts or regulatory agencies charged with enforcing comparable worth policies could face situations requiring choices between competing versions of fair pay structures.

Why do these results so strongly indicate method divergence, when previous investigators often concluded that method differences were negligible? This apparent conflict is not easy to reconcile, in part because previous studies lacked detailed information. However, a number of explanations are possible. First, limitations of the present study might account for the extent of differences among methods. We have already noted shortcomings in the data collection procedures. Furthermore, some method differences could be due to statistical artifacts. The relatively small criterion sample used to develop factor and dimension weights in methods 3 and 4 and the multicollinearity of the traditional plan could have resulted in weighting schemes that were highly sample-specific. However, these shortcomings are not unique to this study. Perhaps more important, they are apt to be typical of current practices in designing and auditing job evaluation plans.

Second, job evaluation plans traditionally have been confined to single occupational families such as factory, technical, clerical, or managerial work, whereas the present study sampled six disparate families. This diversity in the job sample could have magnified differences between methods in compensable factors and weighting schemes.

Third, the earliest comparisons of job evaluation plans (Lawshe, 1945; Lawshe & Maleski, 1946) were between simplified and complete versions of standard point plans similar to methods 1 and 2, and applied to a single job family. Investigators have often cited the high interplan correlations from those studies and have overlooked conflicting evidence from comparisons of distinctly different methods (Atchison & French, 1967). The results obtained here support the intuitive belief that divergence in results will increase with differences in job worth criteria, weighting schemes, and scoring procedures. Finally, the focus in the past typically was on the relative efficiency of alternative methods for predicting or capturing common criterion structures (Robinson, Wahlstrom, & Mecham, 1974). Investigators have interpreted similarity in prediction as agreement between methods and have not investigated effects of method differences on actual decisions about pay levels.

Additional research is needed on all aspects of the development, implementation, and administration of comparable-worth job evaluation plans. However, the most pressing research needs pertain to the basic issue of defining value. The premise of the comparable worth movement is the belief that it is possible to define and measure broadly applicable and acceptable dimensions of job value. Whether such general agreement can prevail across diverse occupational groupings, particularly if historical pay relationships might be disturbed, has yet to be demonstrated. In fact, consensus within occupational families might be difficult to achieve. Employees and employee subgroups differ in how they perceive relevant job-value criteria; comparable worth advocates' frequent references to "men's" and "women's" job-value dimensions serve to demonstrate this fact. More needs to be learned about the extent to which job-value criteria differ, their key determinants and correlates, and whether job evaluation development procedures can modify these values. There is also a need to investigate the extent to which independent job evaluation committees from the same organizations and using the same basic methods, like point plans, generate different factors and weights. Such studies would provide direct and practical tests of how adequate job evaluation procedures are for implementing the principle of comparable worth in wage determination.

In summary, this study investigated the consequences of using different operational definitions of job worth as a basis for pay-related decisions. We developed six different methods of determining job worth and applied them to 206 job classes. Results indicated that choice of evaluation method significantly affects decisions on job hierarchies and pay equity. This finding has major implications for the current debate over the concept of comparable worth.

REFERENCES

AFSCME v. State of Washington, 33 FEP Cases 808 (W.D. Wash. 1983).

American Nurses Association v. State of Illinois, 37 FEP Cases 705 (N.D. Ill. 1985).

Atchison, T., & French, W. 1967. Pay systems for scientists and engineers. ***Industrial Relations***, 7 (1): 44–56.

Arvey, R. D., Passino, E. M., & Lounsbury, J. W. 1977. Job analysis results as influenced by sex of incumbent and sex of analyst. ***Journal of Applied Psychology***, 62: 411–416.

Gomez-Mejia, L. R., Page, R. C., & Tornow, W. W. 1982. A comparison of the practical utility of traditional, statistical, and hybrid job evaluation approaches. ***Academy of Management Journal***, 25: 790–809.

Henderson, R. I. 1985. ***Compensation management: Rewarding performance***. Reston, Va.: Reston Publishing Company, Inc.

Lawshe, C. H., Jr. 1945. Studies in job evaluation II: The adequacy of abbreviated point ratings for hourly paid jobs in three industrial plants. ***Journal of Applied Psychology***, 29: 177–184.

Lawshe, C. H., Jr., & Maleski, A. 1946. Studies in job evaluation III: An analysis of point ratings for salary paid jobs in an industrial plant. ***Journal of Applied Psychology***, 30: 117–128.

Lawshe C. H., Jr., Dudek, E. E., & Wilson, R. F. 1948. Studies in job evaluation VII: A factor analysis of two point-rating methods of job evaluation. ***Journal of Applied Psychology***, 32: 118–129.

Livernash, E. R. (Ed.) 1980. ***Comparable worth: Issues and alternatives***. Washington, D. C.: Equal Employment Advisory Council.

Lorber, L. Z., Kirk, J. R., Samuels, S. L., & Spellman, D. J. 1985. ***Sex and salary***. Alexandria, Va.: The ASPA Foundation.

Madigan, R. M. 1985. Comparable worth judgments: A measurement properties analysis. ***Journal of Applied Psychology***, 70: 137–147.

Mahoney, T. A. 1983. Approaches to the definition of comparable worth. ***Academy of Management Review***, 8: 14–22.

Mayer, C. E. 1984. Equity: Comparable worth debate intensifies. ***Washington Post***, July 22: F1.

McCormick, E. J. 1979. ***Job analysis: Methods and applications***. New York: AMACOM.

McCormick, E. J., Mecham, R. C., & Jeanneret, P. R. 1977. ***Position analysis questionnaire, technical manual***. Logan Utah: PAQ Services, Inc.

Mecham, R. C. 1983. Quantitative job evaluation using the position analysis questionnaire. ***Personnel Administrator***, 28 (6): 82–88, 124.

Nelson, T. 1980. Wage discrimination and comparable worth theory in perspective. ***University of Michigan Journal of Law Reform***, 13 (2): 231–301.

Norwood, J. A. 1982. ***The female-male earnings gap: A review of employment and earnings issues***. Bureau of Labor Statistics Report No. 679. Washington, D.C.: U. S. Government Printing Office.

Nunnally, J. C. 1967. ***Psychometric theory***. New York: McGraw-Hill Book Co.

Ramsey, G. A. 1979. A generalized multiple regression model for predicting college faculty salaries and estimating sex bias. In T. R. Pezzullo & B. E. Brittingham (Eds.), ***Salary equity***: 37–53. Lexington, Mass.: Lexington Books.

Remick, H. 1984. Major issues in a priori applications. In H. Remick (Ed.), ***Comparable worth and wage discrimination***: 99–117. Philadelphia: Temple University Press.

Roberts, H. V. 1980. Statistical biases in the measurement of employment discrimination. In E. R. Livernash (Ed.), ***Comparable worth: Issues and alternatives***: 173–196. Washington, D.C.: Equal Employment Advisory Council.

Robinson, D. D., Wahlstrom, O. W., & Mecham, R. C. 1974. Comparison of job evaluation methods: A policy-capturing approach using the position analysis questionnaire. ***Journal of Applied Psychology***, 59: 633–637.

Sargent, H. A. 1972. Using the point method to measure jobs. In M. Rock (Ed.), ***Handbook of wage and salary administration***: 2/31–2/41. New York: McGraw-Hill Book Co.

Schwab D. P. 1980. Job evaluation and pay setting: Concepts and practices. In E. R. Livernash (Ed.), ***Comparable worth: Issues and alternatives***: 49–78. Washington, D. C.: Equal Employment Advisory Council.

Snelgar, R. J. 1983. The comparability of job evaluation methods in supplying approximately similar classification in rating one job series. ***Personnel Psychology***, 36: 371–380.

Stieber, J. 1959. ***The steel industry wage structure***. Cambridge, Mass.: Harvard University Press.

Treiman, D. J. 1979. ***Job evaluation: An analytic review***. Washington, D.C.: National Academy of Sciences.

Treiman, D. J., & Hartman, H. I. (Eds.) 1981. ***Women, work, and wages: Equal pay for jobs of equal value***. Washington, D. C.: National Academy Press.

Robert M. Madigan earned his Ph.D. degree at Michigan State University. He is an assistant professor of management at Virginia Polytechnic Institute and State University and was formerly a personnel executive in the private sector. His current research focuses on pay equity issues and employment procedures based on validity generalization.

David J. Hoover is an assistant professor of management at Miami (Ohio) University and a Ph.D. candidate in human resource management at Virginia Polytechnic Institute and State University.

© Academy of Management Journal
1986, Vol. 29, No. 1, 101–114.

THE RELATIONSHIP OF SENIORITY TO JOB PERFORMANCE FOLLOWING REINSTATEMENT

ROBERT C. RODGERS
I. B. HELBURN
University of Texas at Austin
JOHN E. HUNTER
Michigan State University

In a seminal study, Ross (1957) found that high seniority employees who were discharged but subsequently reinstated by arbitrators were significantly better performers than reinstated low seniority employees. Only one study since Ross's has replicated this result. Others have reported contradictory evidence. This study proposes a new theory to explain the relationship between seniority and job performance following reinstatement. By quantitatively combining results across studies, we show that senior grievants whom arbitrators exonerated were, just as Ross found, better performers than junior grievants. However, reinstated senior grievants who were not exonerated were found to be significantly poorer performers than nonexonerated junior grievants.

Arbitrators often take seniority into consideration as a mitigating factor when deciding whether to sustain a discharge for just cause or to reinstate a grievant (Elkouri & Elkouri, 1973). Ross (1964: 148–149) argued that cases involving senior employees with good work records should receive special consideration because their infractions were deviations in behavior that are unlikely to be repeated. He based his argument on evidence (Ross, 1957) that reinstated employees with long seniority are more likely to be satisfactory workers than reinstated workers with short seniority.

Ross's research has had widespread influence on arbitration decisions (Jennings & Walters, 1976) even though subsequent findings have been mixed. Of the studies that have examined the relationship between seniority and performance after reinstatement, only one has replicated Ross's original results (McDermott & Newhams, 1971: 531). The bulk of the research on this issue has questioned the presumption that the performance of senior workers after reinstatement is more satisfactory than that of junior workers (Adams, 1978; Jones, 1961; Malinowski, 1981; Rodgers, Helburn, & Hunter, 1983). Jones argued that: "Seniority. . . has little bearing upon whether an employee remains or becomes a satisfactory employee" (1961: 40). On the other hand, Adams found that "reinstated employees with greatest seniority tended to attract more subsequent discipline and discharges" (1978: 74).

The purpose of this study was to reconsider the relationship between seniority and performance after reinstatement in the light of this seemingly contradictory evidence. First, however, we consider the possibility that Ross was correct all along and that the discrepant results found in subsequent studies were simply artifacts of sampling errors.

Obtaining data on the performance of employees reinstated by arbitrators has traditionally been difficult, resulting in a small number of studies with small samples. The number of reinstated workers included in these studies ranges from a low of 16 (Rodgers et al., 1983) to a high of 93 (Adams, 1978). A restriction in range of seniority is also evident. For example, Jones (1961) included no observations for grievants having less than 3 or more than 10 years seniority. Given small sample sizes and dramatic restrictions in range, failure to identify marked patterns between seniority and job performance after reinstatement in these studies would not be surprising.

Populations studied provide another clue as to why previous studies yielded considerably different results. Some of the studies combined exonerated grievants who had been reinstated with full back pay and allowances, including restoration of seniority, with exonerated grievants who forfeited partial or full back pay but had nonetheless been reinstated. Other studies excluded exonerated grievants from their samples either inadvertently or by design. If the relationship between seniority and performance is different for exonerated and nonexonerated employees, mingling the two groups could have generated the equivocal results that have been reported. Thus, in the following section we consider three theoretically plausible models that disentangle exonerated and nonexonerated grievants.

THE MODELS

Model 1

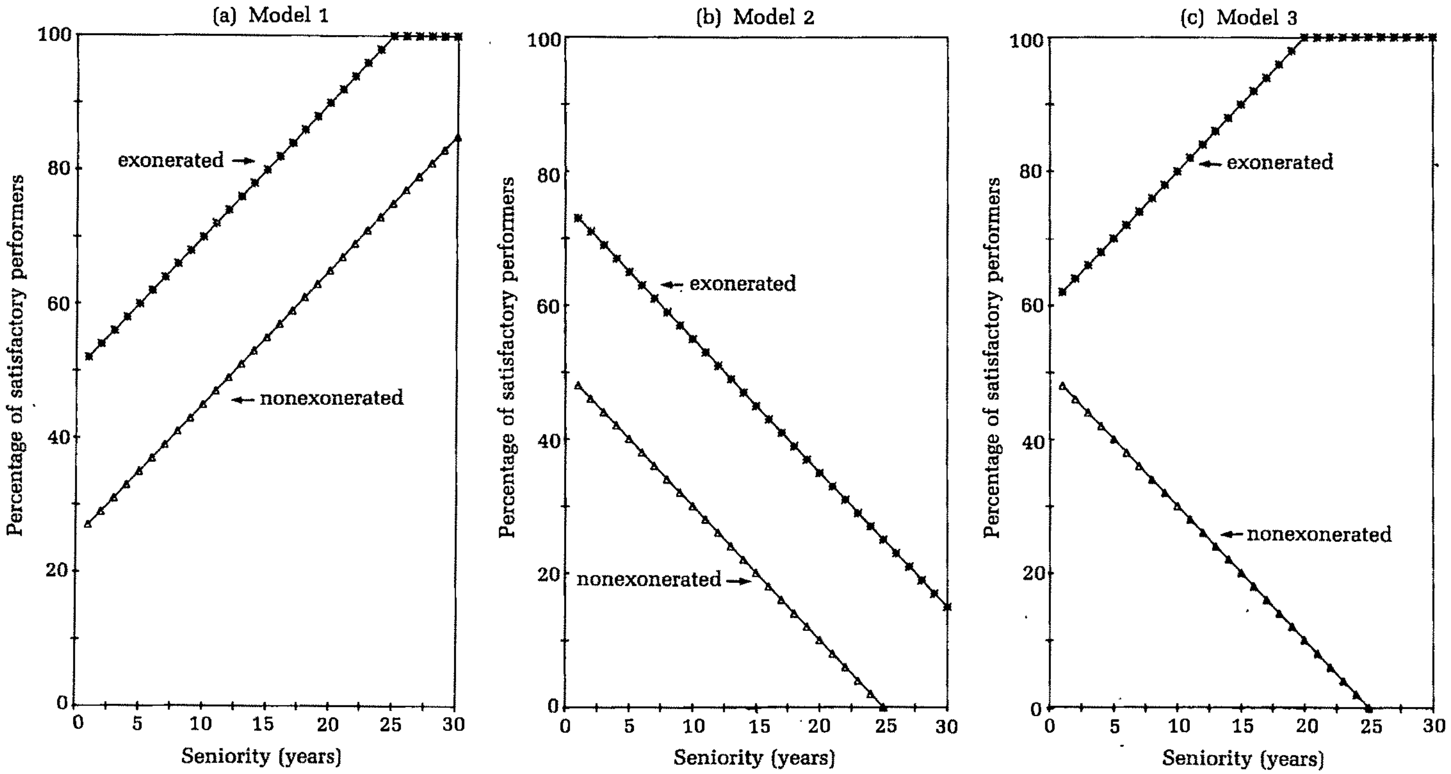
Model 1 suggests that the performance of exonerated workers is better on the average than the performance of nonexonerated workers. And it presumes Ross's conclusion to be valid: after reinstatement, senior workers are in fact better performers than junior workers.

Empirical evidence has shown that the relationship between work experience and job performance for the working population in general is positive. The average correlation between job experience and work performance across 425 studies has been shown to be $r = +.18$ (Hunter & Hunter, 1984). Thus, whether a worker has been exonerated or not, on the average, senior workers' performance after reinstatement should be better than junior workers' performance after reinstatement. Figure 1a depicts the positive regression slopes for exonerated and nonexonerated workers implied by this model.

Model 2

Model 2 assumes that grievants with high seniority will perform poorly after reinstatement whether or not they were exonerated. Perhaps employers always conduct systematic reviews of grievants' work histories before

FIGURE 1
Models of the Relationship between Seniority and the Performance of Exonerated and Nonexonerated Grievants Following Reinstatement



discharging them, regardless of the issues involved or the severity of infractions. We would expect employers to require flagrant and repeated demonstrations of unsatisfactory performance before discharging any senior employee. Even employers who do not care about justice or loyalty might fear that unjust discharges of high performing senior workers will chill the motivation of workers who remain.

Generally speaking, arbitrators are also somewhat disposed to give senior workers second chances, regardless of the nature of infractions (Ross, 1964: 148-149). Thus, because companies and arbitrators apply more lenient standards to senior workers than to junior workers, we might expect the job performance of senior workers to be poorer than that of junior workers after reinstatement. The performance of exonerated employees should still be better than the performance of nonexonerated workers, however (see Figure 1b).

Model 3

The first two models ignore the nature of infractions. Model 3 contrasts two different situations: repeated minor offenses versus one-time major offenses. For example, consider repeated tardiness or repeated absenteeism. In these instances, it is likely that employers would consider work records before deciding to discharge employees. Newly hired individuals are likely to be fired regardless of work records, but long-term employees are likely to be fired only if they have bad work records. If this is true, there would be a negative correlation between seniority and performance among those fired for repeated minor infractions. In this case, it is also likely that guilt is not the issue; rather, the issue is whether a worker deserves a second chance or not. Workers who get second chances are reinstated but not exonerated.

On the other side might be allegations of theft or heated arguments with supervisors. Here guilt would be a primary issue and work record would not enter into consideration. Such workers would be reinstated only if exonerated. In these cases, accused workers are likely to constitute a random sample in regard to performance. Thus, we would predict the usual positive correlation between experience and performance.

Model 3 thus predicts that the regression lines for exonerated and nonexonerated workers will have opposite slopes: performance will be higher for senior workers among those exonerated and performance will be lower for senior workers among those who were not exonerated. Figure 1c shows these predicted relationships.

A direct test of model 3 would require data relating numbers, types, and severity of infractions. Unfortunately, none of the previous studies reported these data. In addition, there is no way to ascertain from the studies how carefully employers reviewed employee work histories before rendering discharge decisions. The present analysis used whether grievants were exonerated or not as a proxy for the severity and frequency of infractions leading to discharges. Model 3 predicts that workers fired for repeated minor infractions will rarely be exonerated; it also predicts that workers accused of major infractions will be reinstated only if exonerated. Thus, model 3 assumes a high correlation between whether workers were exonerated and whether infractions were one-shot major offenses or repeated minor offenses.

Separate data are not available for exonerated and nonexonerated workers. Rather, some previous studies considered only nonexonerated workers and other studies included both groups. It is mathematically possible to obtain estimates for exonerated workers by using these two sets of study results, as the Appendix shows.

METHODS

A search of the arbitration literature was conducted in journals specializing in labor and industrial relations issues, including *Industrial and Labor Relations Review, Industrial Relations, Journal of Labor Research, National Academy of Arbitrators Proceedings, Arbitration Journal,* and the *Industrial Relations Research Association Proceedings.* We located six relevant studies.

To cumulate results quantitatively, results from each study must be converted to a common metric (Hunter, Schmidt, & Jackson, 1982). The data found in the located studies were converted to two metrics: (1) the proportion of satisfactory employees within designated seniority intervals, and (2) the correlation coefficient.

Table 1 summarizes data reported for each study on the job performance of 219 grievants who were reinstated by arbitrators. Cumulative estimates that include an additional 58 grievants from the Adams (1978) study appear at the bottom of Table 1, increasing the total sample size to 277.

Cumulative estimates for the 219 grievants show that 47 percent of the workers with two years or less seniority performed satisfactorily after reinstatement, and 53 percent of those with 20 or more years performed satisfactorily. When the intervals are collapsed, enabling inclusion of all of Adams's (1978) results, the trend across seniority categories, which were 0-5, 6-10, 11-20, and over 20 years, is insignificant ($\chi^2 = 4.313$, $df = 3$).

Since the intervals have different ranges and different weights, these comparisons are based on interval estimates of varying degrees of reliability. For example, a much larger number of reinstated workers fell within the lowest seniority interval than within any other interval. The prediction that 54 percent of 145 reinstated workers with 0-5 years seniority will perform satisfactorily upon reinstatement clearly has considerably less sampling error than an estimate that 53 percent of 15 employees with over 20 years seniority will have satisfactory performance.

Correlations are a more useful statistic for considering whether differences in estimates across studies are due to sampling error or to true differences in study populations. We drew correlations from the studies as follows: Rodgers and colleagues (1983) reported a correlation between seniority and job performance after reinstatement of $r = -.25$. Jones (1961) reported seniority and job performance data for each reinstated grievant. Using these data, we correlated seniority with performance. Performance after reinstatement

TABLE 1
Performance Following Reinstatement by Seniority

	Years of Seniority at Time of Discharge												
	0-2		3-5		6-10		11-15		16-20		> 20		
Studies	S[a]	U[a]	S	U	S	U	S	U	S	U	S	U	Totals
Ross (1957)	7	11	13	5	14	2	6	0	3	0	0	1	62
Jones (1961)	0	0	7	5	2	3	0	0	0	0	0	0	17
McDermott & Newhams (1971)	3	5	7	7	7	1	6	1	4	1	3	2	47
Adams[b] (1978)	n.r.[c]		n.r.		17	12	n.r.		n.r.		2	4	35
Malinowski (1981)	0	6	0	1	7	4	2	4	1	0	2	0	42
Rodgers et al. (1983)	2	2	1	2	5	3	0	0	0	0	1	0	16
Totals	21	24	34	20	52	25	14	5	8	1	8	7	219
Proportion satisfactory	.47		.63		.68		.74		.89		.53		

	Cumulative Estimates Using Collapsed Intervals, All Studies								
	0-5		6-10		11-20		> 20		Total
	S	U	S	U	S	U	S	U	
Totals	79	66	52	25	26	14	8	7	277
Proportion satisfactory	.54		.68		.65		.53		

[a] S = number of satisfactory performers; U = number of unsatisfactory performers.

[b] Adams used a 0-5 and a 11-20 year interval. All of Adams's data are included in the summary at the bottom of the table.

[c] n.r. = not reported for this seniority interval.

was measured by 1 = above average; 0 = average; -1 = below average. This yielded a correlation of r = -.38. Using the data in Table 1, we approximated correlations in the other four studies. Seniority for each worker was estimated using the mean of the interval midpoint for that worker. Performance ratings were successful =1 and unsuccessful =0. We did not include data reported in the over-20 interval in the tabulation because the interval midpoint could not be determined.

Table 2 reports the correlations (r's) that were calculated using these procedures. Had the studies reported the raw data with which to calculate the correlations, the coefficients could have been slightly different. With as many as five intervals, however, the difference between the correlations estimated from interval data and the correlations that might have been calculated from individual observations is trivial. Although correlations calculated from individual observations would still have been preferable, study correlations calculated from interval data are close estimates of those correlations.

Observed correlations were compared to see whether discrepancies across studies were due to true differences in study populations or to sampling error. The procedure had five steps. First we calculated a weighted average correlation. Second, we calculated the observed weighted variance in the correlations. Third, we estimated the proportion of the variance due to sampling error using the standard formula. Fourth, we subtracted the proportion of the variance due to sampling error from the observed variance, providing our best estimate of the population variance across studies, with sampling error taken into account. Fifth, since our estimate of the population variance was not zero and therefore could not be explained by sampling error, we tested the moderator hypotheses by grouping the studies into two subsets. The first subset included studies with grievants who were reinstated with full back pay and grievants who were reinstated with less-than-full back pay—a mixed subset of exonerated and nonexonerated individuals. The second subset included studies of grievants who were reinstated with partial or no back pay—nonexonerated individuals. We then repeated the first four procedures for each subset. If the population variance was zero for both subsets, we could conclude that performance was moderated by inclusion of exonerated grievants in a study sample.

RESULTS USING CORRELATIONS AS THE COMMON METRIC

First, assuming that the true population correlation is constant across studies, our best estimate of the population correlation r was a weighted average, where N_i is the number of reinstated workers in study i:

$$r = \frac{\Sigma(N_i r_i)}{\Sigma N_i} = .03 \quad .$$

TABLE 2
Characteristics of Studies

Studies	*n*	*r*	Seniority		Performance		Proportion Exonerated
			Means	σ	Means	σ	
Mixed subset[a]							
Malinowski (1981)	42	.04	6.59	6.23	.64	.49	.39
McDermott & Newhams (1971)	47	.21	9.23	7.61	.64	.49	.35
Ross (1957)	62	.29	6.05	5.23	.65	.47	.30
Total	151						
Subset means (weighted)		.20	7.19	6.48	.64	.48	.34
Nonexonerated subset							
Adams (1978)	93	-.13	7.12	6.29	.51	.50	.00
Jones (1961)	17	-.38	5.75	3.46	.53	.51	.00
Rodgers et al. (1983)	16	-.25	7.08	6.38	.44	.51	.00
Total	126						
Subset means (weighted)		-.18	6.93	6.02	.50	.50	.00

[a] Subset consisted of studies with both exonerated and nonexonerated employees.

As would be expected from the generally inconsistent findings summarized in Tables 1 and 2, the average correlation between seniority and performance after reinstatement is quite small.

Second, we calculated the variance of r across studies. The best estimate of the variance across studies was the frequency-weighted average squared error, where r = the mean weighted correlation (.03):

$$\sigma_r^2 = \frac{\Sigma [N_i\ (r_i - r)^2]}{\Sigma N_i} = .0441\ ,$$

And $\sigma_r = .21$. Third, a portion of this variance may still be due to sampling error σ_e^2. The magnitude of this error, where k = the number of studies, was estimated as follows:

$$\sigma_e^2 = \frac{(1 - r^2)^2 k}{\Sigma N_i} = .0216\ .$$

Fourth, the population variance σ_ρ^2 was obtained by subtracting the variance due to sampling error from the observed variance:

$$\sigma_\rho^2 = \sigma_r^2 - \sigma_e^2 = .0441 - .0216 = .0225\ .$$

Thus, taking sampling error into account dropped our estimate of the standard deviation of the population correlation from $\sigma_r = .21$ to $\sigma_\rho = .15$ (the square root of .0225). However, this is still a large variation. A formal statistical test for determining whether no variation exists (Hunter et al., 1982: 45-47) showed that the variation across studies was greater than what might be dictated by chance alone ($\chi^2 = 12.24$, $df = 5$). This analysis suggested the presence of a moderator variable, as model 3 predicted.

The reinstatement studies were thus divided into two subsets, those with only workers who had not been exonerated and those with mixed samples. Three studies (Adams, 1978; Jones, 1961; Rodgers et al., 1983) included no employees who had been exonerated by arbitrators (see Table 2). These studies formed one subset. The three remaining studies (Malinowski, 1981; McDermott & Newhams, 1971; Ross, 1957), which included 39, 35, and 30 percent exonerated employees respectively, formed a second, mixed subset. Two study statistics were calculated for each of the subsets—the mean correlation and the corrected variance. The first subset consisted of studies that included only employees who had not been exonerated ($n = 126$). The second subset included studies that mixed exonerated employees with nonexonerated employees ($n = 151$). The correlation is negative ($r = -.18, n = 126$) for the nonexonerated group and positive ($r = +.20$, $n = 151$) for the mixed group. The difference between the correlations is also significant at the .001 level ($z = 3.13$).

TABLE 3
Analysis of the Subsets

Subset Statistics	Nonexonerated	Mixed[a]
r	-.18	+.20
σ^2_r	.0079	.0105
σ^2_e	.0223	.0198
σ^2_p	-.0144	-.0093
σ_p	0	0

[a]Subset consisted of studies with both exonerated and nonexonerated employees.

Analysis of the subset characteristics, shown in Table 3, indicates that including exonerated workers in a study sample moderates the relationship between seniority and performance after reinstatement. This is illustrated by the finding that the weighted correlation differs considerably across the subsets and the corrected variance σ_p is zero in each subset (Hunter et al., 1982: 48). Studies that included no exonerated employees were more likely to find an inverse relationship between seniority and performance. Studies that included exonerated employees in their samples were more likely to find a positive relationship between seniority and performance. This explains why the weighted mean correlation is so small ($r = .03$) when the six studies are combined. We thus have reasonable cause to believe that the regression slope for exonerated workers is not parallel to the regression slope for nonexonerated workers. This tends to disprove models 1 and 2.

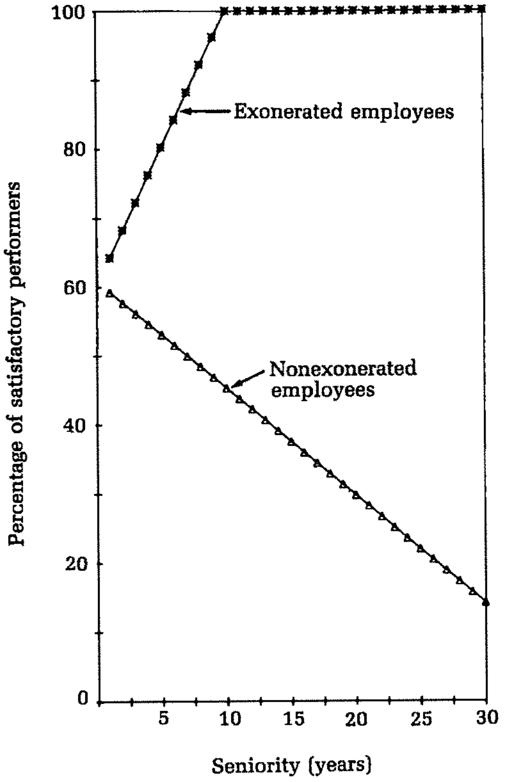
Testing the model 3 hypotheses requires data for two groups of employees, those who are exonerated and those who were not. We have pure data on employees who were not exonerated, but data for exonerated workers are mixed with data on those who were not. We found the strength of the moderated relationships by calculating the regression slope for the known nonexonerated group and deriving the regression slope for the unknown exonerated group. As the Appendix reports, the regression slope for the exonerated group is $b = .040$; for the group that was not exonerated, it was $b = -.015$. Thus, these calculations provide support for model 3.

Figure 2 illustrates regression slopes based on the derived estimates between seniority before discharge and job performance following reinstatement for exonerated workers and for workers who were not exonerated. Exonerated workers with 10 years or more seniority almost always perform satisfactorily after reinstatement. A steep decline appears for workers who were not exonerated, suggesting that less than 20 percent with 30 years seniority perform satisfactorily.

DISCUSSION

For a topic such as ours, where previous research is scant and empirical data quite difficult to obtain, the quantitative cumulation of results can be

FIGURE 2
Predicted Relationship between the Performance of Exonerated and Nonexonerated Grievants Following Reinstatement



useful. Our analysis has indicated that the apparently contradictory evidence in the literature on seniority and job performance following reinstatement should not be interpreted as inconclusive. Rather, cumulated results demonstrated support for one of three theoretical models. We found a positive relationship between seniority and job performance following reinstatement for grievants who had been exonerated and a negative relationship for grievants who had not been exonerated by arbitrators.

Our estimate of the correlation between seniority and job performance following reinstatement for exonerated workers of r = + .20 is remarkably similar to Hunter and Hunter's (1984) true population estimate of the correlation between work experience and performance. The results thus vindicate Ross's (1957) original conclusion, but only for workers who were exonerated.

Model 3 implies that the critical factor is not seniority per se; it is probably the type of infraction that triggers a discharge. When one time major infractions are involved, companies are unlikely to review employee work records. We speculate that with multiple infractions companies are likely to review work histories and, as a result, to take seniority into consideration as a mitigating factor. This results in the discharge of senior workers with unusually poor work histories.

Because of data limitations, our test of model 3 has been indirect. We could extract no data from the studies regarding the nature or the frequency of previous infractions. We hope that future studies will take these factors into consideration.

The practical implication of our analysis is apparent. In certain instances—such as when multiple infractions trigger a discharge—we suspect that a review of a grievant's work history is part of the discharge process. Under such circumstances, arbitrators should not give special consideration to grievants who have high seniority. Contrary to conventional wisdom, our results indicate that their performance after reinstatement will be less satisfactory than the performance of junior grievants.

REFERENCES

Adams, G. W. 1978. ***Grievance arbitration of discharge cases.*** Kingston, Ontario: Industrial Relations Centre, Queen's University at Kingston.

Elkouri, F., & Elkouri, E. A. 1973. ***How arbitration works*** (3rd ed.). Washington, D. C.: Bureau of National Affairs.

Hunter, J. E., & Hunter, R. F. 1984. The validity and utility of alternative predictors of job performance. ***Psychological Bulletin,*** 96(1): 72–98.

Hunter, J. E., Schmidt, F. L., & Jackson, G. B. 1982. ***Meta-analysis: Cumulating research findings across studies.*** Beverly Hills, Calif.: Sage Publications.

Jennings, K., & Walters, R. 1976. Discharge cases reconsidered. ***The Arbitration Journal,*** 31: 164–180.

Jones, D. L. 1961. ***Arbitration and industrial discipline.*** Ann Arbor, Mich.: Bureau of Industrial Relations.

Malinowski, A. A. 1981. An empirical analysis of discharge cases and the work history of employees reinstated by arbitrators. ***The Arbitration Journal,*** 36: 31–46.

McDermott, T. J., & Newhams, T. H. 1971. Discharge-reinstatement: What happens thereafter? ***Industrial and Labor Relations Review,*** 24: 526–540.

Rodgers, R. C., Helburn, I. B., & Hunter, J. E. 1983. ***Cumulating results across studies: Is seniority related to performance?*** Working paper No. 83/84-4-6, College of Business Administration, University of Texas, Austin.

Ross, A. M. 1957. The arbitration of discharge cases: What happens after reinstatement? ***Proceedings of the tenth annual meeting of the National Academy of Arbitrators***: 21–60. Washington, D. C.: Bureau of National Affairs.

Ross, A. M. 1964. Discussion on the criminal law and industrial systems. ***Labor arbitration—perspective and problems: Proceedings of the seventeenth annual meeting, National Academy of Arbitrators***: 144–152. Washington, D. C.: Bureau of National Affairs.

APPENDIX

Derivation of the Regression Slopes

The slope of the regression coefficient for the nonexonerated subset is calculated as follows:

$$b_{\text{nonexon}} = r \times \left(\frac{\sigma_{\text{perf}}}{\sigma_{\text{sen}}}\right) ,$$

$$b_{\text{nonexon}} = -.18 \times \left(\frac{.50}{6.02}\right) = -.015 .$$

The slope of the regression for the exonerated subset must be derived. The mean covariances and the mean correlations for the exonerated group and the nonexonerated group affect the regression line derived from the mixed study data. The variance for seniority and performance in the exonerated group is derived from the fact that the variance of the mixed group is equal to the variance of the means plus the mean of the within group variances. Characteristics of the exonerated group are not known, but may be derived from the fact that the mean ($\bar{x}$) of the mixed group is equal to the weighted mean ($\bar{x}_{\text{exon}}$) for the exonerated group plus the weighted mean ($\bar{x}_{\text{nonexon}}$) for the nonexonerated group. In the mixed studies exonerated workers constituted 34 percent of the population, such that

$$\bar{x}_{\text{mixed}} = .34\,\bar{x}_{\text{exon}} + .66\,\bar{x}_{\text{nonexon}} .$$

Calculation of mean seniority for the exonerated group $\bar{x}_{\text{exon}}$ is thus:

$$\bar{x}_{\text{exon}} = \frac{7.19 - .66 \times 6.93}{.34} = 7.70 .$$

Similarly, mean performance for the exonerated group $\bar{P}_{\text{exon}}$ is:

$$\bar{P}_{\text{exon}} = \frac{.64 - .66 \times .50}{.34} = .91 .$$

For the mixed group,

$$\sigma^2 = \sigma^2_{\text{means}} + \text{average}\,(\sigma^2_{\text{within}}) .$$

Taking the observed mean on seniority for the nonexonerated from Table 2 (6.93) and the derived mean on seniority for the exonerated (7.70), the σ_{means} is:

$$.66 \times (6.93 - 7.19)^2 + .34 \times (7.70 - 7.19)^2 = .13 .$$

Further, where σ^2 on seniority for the nonexonerated $= 6.02^2 = 36.24$ and the σ^2 on seniority for the mixed group $= 6.48^2 = 41.99$, we can solve for the unknown V, the σ^2 for the exonerated as follows:

$$41.99 = .13 + [.66 \times (36.24) + .34V] ,$$

$$V = 52.77,$$

and thus $\sigma = 7.26$. The calculation of the standard deviation of performance for the exonerated group is:

$$\sigma^2_{\text{exon}} = .91 \times .09 \,,$$

$$\sigma_{\text{exon}} = .29 \,.$$

The correlation between seniority and performance for the exonerated group is readily obtained from the fact that the covariance of the mixed subset is equal to the covariance of the means plus the mean covariances.

The covariance of the nonexonerated is $r_{\text{nonexon}} \times \sigma_{\text{perf}} \times \sigma_{\text{sen}}$ or $6.02 \times .50 \times -.18 = -.542$; the covariance of the mixed group is $6.48 \times .48 \times .20 = .622$. And, the covariance of the means is equal to: $.66 \times [(7.19 - 6.93) \times (.64 - .50)] + .34 \times [(7.19 - 7.70) \times (.91 - .64)] = .07$. Solving for the unknown covariance (C) of the exonerated group,

$$.622 = .07 + [(.66 \times -.542) + .34\ C] \,,$$

$$C = 2.67 \,,$$

$$r = \frac{2.67}{7.26 \times .29} \,.$$

For the exonerated group, our best estimate is thus that $r = 1.00$. Therefore, with $r = 1.00$, a σ for performance of .29, and σ for seniority of 7.26, the regression slope b is:

$$b = 1.00 \times \frac{.29}{7.26} = .040 \,.$$

Robert C. Rodgers earned his Ph.D. degree from Michigan State University; he is Assistant Professor of Management at the Graduate School of Business, University of Texas at Austin. His current research interests include organizational designs for public sector organizations, administrative policies for public sector managers, human resource management, and research methods.

I. B. Helburn earned his Ph.D. degree from the University of Wisconsin, Madison; he is the Bobbie and Coulter R. Sublett Centennial Professor, Graduate School of Business, University of Texas, Austin. His current research interests include labor relations, with specific emphasis on profit sharing and arbitration decision making.

John E. Hunter earned his Ph.D. degree from the University of Illinois; he is Professor of Psychology, Michigan State University. His current research interests include personnel selection, training and experience, and the relationship between knowledge and performance.

© Academy of Management Journal
1986, Vol. 29, No. 1, 115–129.

EFFECTS OF GENDER ON SELF- AND SUPERVISORY RATINGS

LYNN McFARLANE SHORE
GEORGE C. THORNTON III
Colorado State University

This research investigated the effects of supervisors' and subordinates' genders on self- and supervisory ratings in an organizational setting. Participants were assemblers, 35 men and 35 women, and their supervisors, 16 men and 19 women. Results showed that subordinates' self-ratings were higher than their supervisors' ratings of them and that gender did not affect the relationship between self- and supervisory ratings. In prior research involving novel situations in laboratory settings, women provided lower self-ratings and higher ratings of others than men. The present study showed no gender differences in ratings on familiar tasks in a real work setting in which performance feedback was available.

Concern for minimizing error in ratings has produced many and varied approaches to investigating accuracy in performance appraisals (Landy & Farr, 1980). The present study synthesized two distinct research approaches to accuracy and fairness in performance evaluations. First, much research has compared self- and supervisory ratings in work settings in an effort to evaluate the accuracy of self-assessments for organizational purposes (Thornton, 1980). Second, concern about fair evaluations of men and women has led to extensive laboratory research on gender effects in evaluation situations (Nieva & Gutek, 1980). The present study integrated these two areas of research to address issues of accuracy raised by both literatures.

Theory and selected research results suggest that previous laboratory findings about gender differences in ratings may have limited generalizability to the workplace. Therefore, the present study investigated gender differences in a realistic setting. Men and women employees and supervisors in an actual work organization participated in the study.

REVIEW OF LITERATURE AND HYPOTHESES

Congruence between Self- and Supervisory Ratings

Many researchers have suggested that self-ratings serve a variety of organizational purposes (Meyer, 1977; Primoff, 1980). The use of self-ratings has

---

We gratefully acknowledge the assistance of James McCambridge for his contributions to the development of this research project; Alan Inada for his help in data collection; and Ted H. Shore for his editorial assistance.

been found to decrease subordinate defensiveness in performance appraisal interviews (Bassett & Meyer, 1968) and to increase accuracy in evaluations (Bartlett, 1959). Self-ratings also serve as criteria against which to validate other assessment procedures, thereby helping to make criterion-related validity studies feasible in organizations (Levine, Flory, & Ash, 1977). However, concern about the use of self-ratings in organizations has arisen because extensive research has shown that self-ratings tend to be more lenient than supervisory ratings (Thornton, 1980). Leniency of self-ratings relative to supervisory ratings has been found for clerical workers (Parker, Taylor, Barrett, & Martens, 1959), technical employees (Kirchner, 1966), nurses (Klimoski & London, 1974), first-level supervisors (Waldman & Thornton, 1979), and executives (Holzbach, 1978; Thornton, 1968).

Although most research shows that self-ratings are more lenient than supervisory ratings, there have been exceptions (Heneman, 1974; Waldman, 1979). In fact, several variables appear to influence leniency in self-ratings. These include tenure (Brief, Aldag, & VanSell, 1977), age (Lawler, 1967), and self-esteem (Baird, 1977). One variable that comparisons of self- and supervisory ratings have not investigated is gender, even though extensive research has shown gender effects in evaluation situations (Nieva & Gutek, 1980). Thus, the purpose of the present study was to evaluate the effects of supervisors' and subordinates' genders on the congruence of self- and supervisory ratings.

Gender of Subordinates

Waldman (1979) suggested that the gender of subordinates may affect the leniency of self-ratings. Despite the well-documented importance of gender as a moderator in many situations (Maccoby & Jacklin, 1974), most studies comparing self- and supervisory ratings do not state whether the subjects were men or women. However, Dipboye and Flanagan (1979) reviewed all the empirical articles from the 1966, 1970, and 1974 volumes of *Journal of Applied Psychology*, *Organizational Behavior and Human Performance*, and *Personnel Psychology*, and found that field research in industrial-organizational psychology has predominately had men as subjects. Thus, although most studies have shown that subordinates tend to rate themselves higher than their supervisors rate them, this finding may only apply to men. In fact, Waldman (1979) found that self-ratings were lower than supervisory ratings when a group of women secretaries rated themselves. Since the results of that study came entirely from a sample of women subordinates, it did not conclusively determine the importance of subordinate gender in comparisons of self- and supervisory ratings. However, Waldman's study did lend support to the notion that self-ratings may be more lenient than supervisory ratings for men subordinates but not for women subordinates.

In a comprehensive review of the literature on gender differences in attitudes and behavior, Maccoby and Jacklin (1974) concluded that women have less positive self-referent attitudes than men in many types of situations. Specifically, women were found to manifest lower performance expectancies, self-evaluations of abilities, and self-evaluations of completed performances. Lenney's (1977) literature review of self-confidence in achievement settings for men and women drew conclusions that were consistent with Maccoby and Jacklin's. Lenney theorized that in many situations, especially where feedback is absent or where social cues are salient, men tend to overestimate their performance and women tend to underestimate their performance. Therefore, certain situational variables, such as feedback, which exist in real work settings but not in laboratories may influence self-ratings. A subsequent study by Lenney, Browning, and Mitchell (1980) showed that women underestimated their performance more than men did when performance criteria were ambiguous but not when performance criteria were clear. Finally, Deaux and Farris (1977) found that men evaluated their performance more favorably than did women, despite equivalent objective scores.

All of the literature cited above on gender differences in self-ratings is based on laboratory research. We found only two studies that used actual employees in organizational settings. Deaux's study (1979) of managers was consistent with the laboratory research in its finding that men not only evaluate their performance more favorably than women, but also rate themselves as having more ability and greater intelligence. In contrast, Snyder and Bruning (1979) did not find gender differences in self-rated ability to perform a job effectively. Overall, most prior research on gender differences in self-evaluations suggests that women rate themselves lower than men do. However, since most of this research took place in laboratory settings, this conclusion may not generalize to research conducted in work environments where situational variables like feedback may play important roles.

The gender of subordinates may also influence supervisory ratings. However, the literature in this area is not consistent. Some studies have shown that people evaluate men more highly than women (Fidell, 1970; Pheterson, Kiesler, & Goldberg, 1971; Rosen & Jerdee, 1974; Schmitt & Lappin, 1980); other researchers have shown bias favoring women in evaluations (Bigoness, 1976; Hamner, Kim, Baird, & Bigoness, 1974; Jacobson & Effertz, 1974; Mobley, 1982; Taynor & Deaux, 1973); and still others have shown no differences in evaluations of men and women (Frank & Drucker, 1977; Hall & Hall, 1976).

After an extensive review of the literature on gender effects in evaluation, Nieva and Gutek (1980) theorized that the inconsistency in the literature may be due to the varying amounts of information provided to raters regarding the performance levels of ratees. Specifically, they suggested that raters tend to rely more on stereotypes in making ratings when they have little information on the actual performance of an individual. Furthermore, Terborg and Ilgen (1975) suggested that stereotypes influence evaluations most when evaluators know little about individuals they are rating. This is often the case in, for example, hiring decisions. They also suggested that the effect of gender role stereotypes diminishes as evaluators obtain more information about workers. In order to test these theories, we carried out the present study in a situation in which raters would not rely so much on stereotypes.

Since the information most supervisors have about their subordinates is considerable, especially compared to what an average subject in a laboratory study has, we expected that the gender of subordinates would not influence supervisory ratings in an actual work setting.

Gender of Supervisors

Brief and colleagues (1977) suggested that investigations of variables that influence the level of agreement between self- and supervisory ratings should consider matches between characteristics of supervisors and subordinates rather than just characteristics of the subordinates. A great deal of literature indicates that raters' genders may influence ratings of others just as subordinates' genders may influence self-ratings.

The majority of studies investigating raters' genders have shown a trend for women to rate others higher than do men. Elmore and LaPointe (1974, 1975) found no differences between men and women students' ratings of teachers, except that women students rated instructors higher on specifying course objectives. Lee and Alvares (1977) did not find differences in men and women undergraduates' ratings of supervisors in a simulated industrial setting. However, women described the same supervisors as being higher in consideration than did men. London and Poplawski (1976) found that undergraduate women in performance appraisal and interview simulations gave higher ratings than undergraduate men. Hamner and colleagues (1974) also found that women gave higher ratings than men in a simulated work sampling task. In sum, the literature shows a general trend for women to give higher ratings than men. On the basis of these findings, we expected that the gender of supervisors may influence comparisons of self- and supervisory ratings.

Increasing numbers of women supervise men as well as women subordinates (Terborg, 1977; Wexley & Pulakos, 1982). As yet, very little research has examined the consequences of this trend. The extent of gender differences in self-ratings and supervisory ratings in work organizations remains largely undocumented, despite the important implications of such differences. Furthermore, the findings of research on gender differences in self- and supervisory ratings are inconsistent and derive mostly from research conducted in laboratory settings. Again, a field study is clearly called for. Although most prior research would indicate that women employees rate themselves lower than their supervisors rate them, enough theoretical foundation exists to suggest that a different set of social dynamics takes place in applied settings in which women are working successfully at familiar work tasks. These dynamics may lead to a different pattern of supervisory and subordinate ratings. In light of the lack of field research and the conflicting evidence from previous studies on gender differences in ratings, the present study investigated the effects of supervisors' and subordinates' genders on self- and supervisory performance ratings.

To summarize, a comprehensive review of the literature suggests that both supervisors' and subordinates' genders may affect levels of agreement between self- and supervisory ratings because: (1) men tend to evaluate their performance more favorably than women rate theirs, indicating potential gender differences in self-ratings by subordinates; and (2) women tend to rate others more favorably than men rate others, showing potential gender differences in supervisory ratings. Therefore, the present study investigated the following hypotheses:

> *Hypothesis 1: Women subordinates will rate themselves lower than their supervisors rate them whereas men subordinates will rate themselves higher than their supervisors rate them.*
>
> *Hypothesis 2: Women subordinates will rate themselves lower than men subordinates will rate themselves.*
>
> *Hypothesis 3: Women supervisors will rate subordinates higher than subordinates rate themselves, whereas men supervisors will rate subordinates lower than subordinates rate themselves.*
>
> *Hypothesis 4: Women supervisors will rate subordinates higher than men supervisors will rate their subordinates.*

METHODS

Participants and Design

The participants were assemblers and their supervisors in a large electronics manufacturing company in the Rocky Mountain region. The assemblers included 35 men and 35 women; the supervisors included 16 men and 19 women. The average age of the assemblers was 35, and their average job tenure was 2.75 years. Only supervisors with both men and women subordinates were asked to participate in the study; 98 percent agreed to do so and each provided performance appraisal information for one man and one woman subordinate. Therefore, supervisors with subordinates of only one gender did not participate in the study. Subordinates were selected at random for participation, which was voluntary for both supervisors and subordinates.

A 2 × 2 × 2 factorial design was used with (1) source of rating, supervisor or self; (2) gender of subordinate; and (3) gender of supervisor as factors. The dependent measure was the rating of subordinate's job performance.

Questionnaire

The questionnaire had two sections, a cover letter and a performance appraisal form. Supervisors' and subordinates' questionnaires had slightly different wordings since subordinates rated themselves and supervisors rated subordinates.

The purpose of the cover letter was to provide supervisors and subordinates with general instructions and information. Each participant's cover letter bore a unique identification number that matched a corresponding master list of names held only by the researchers at Colorado State University. Respondents were told that the purpose of the identification number was to make sure both members of a supervisor-subordinate pair had participated.

They were also told that researchers from the University would contact them if their forms were not received. For supervisors, cover letters also said which of their subordinates to evaluate. All participants were informed that their appraisals would remain strictly confidential and would not affect the regularly scheduled performance planning and review process.

The performance appraisal form was very similar to the form regularly used in the firm for evaluating the performance of assemblers. However, some wording changes were necessary for self-ratings because the company did not use self-ratings. The performance appraisal form had six dimensions: (1) work quality, (2) work quantity, (3) judgment, (4) initiative, (5) teamwork, and (6) dependability. Subordinates were rated on a 1- to 5-point scale, where 1 indicated performance below the minimum acceptable level and 5 indicated that performance consistently far exceeded expectations.

Data Collection

All subordinates and corresponding supervisors received cover letters and performance appraisal forms. Supervisors were instructed to complete an appraisal for each of their subordinates participating in the study. Subordinates were asked to complete self- appraisal forms. Subordinates and supervisors were told to put their completed questionnaires into envelopes provided them, seal the envelopes, and mail them to the personnel department. We contacted those who did not return their questionnaires to urge their participation and got back a total of 98 percent of the questionnaires. Upon receiving the completed questionnaires, we mailed a letter to each of the participants explaining the purpose and results of the study.

RESULTS

Correlations among the six performance dimensions on the performance appraisal form resulted in values ranging from .29 to .60. Since many of these correlations were low, we concluded that it would be inappropriate to combine the six dimensions into an overall index of job performance (James, 1973). Thus, we analyzed the performance dimensions separately for all of the hypotheses.

A split plot factorial design (Kirk, 1982) was used to analyze the results because confounding of groups of supervisors and genders of supervisors was unavoidable even though all treatments were crossed. Supervisor's gender was a between-blocks factor, with some supervisors men and others women, since supervisor's gender was confounded with differences between blocks or groups of supervisors. Subordinate's gender and source of rating were within-blocks factors, and thus not confounded, because (1) each supervisor rated one man and one woman subordinate, and (2) for all subordinates, ratings were obtained from both the individuals and their supervisors.

Table 1 presents mean performance ratings for all conditions. The six performance dimensions were conceptually related to each other since they were all facets of perceived work performance. Therefore, a multivariate analysis of variance was an appropriate first analysis. If the multivariate test

TABLE 1
Mean Performance Ratings by Supervisor's Gender, Subordinate's Gender, and Source of Rating[a]

	Men Supervisors				Women Supervisors			
	Men Subordinates		Women Subordinates		Men Subordinates		Women Subordinates	
Performance Dimensions	Supervisory Ratings	Self-Ratings	Supervisory Ratings	Self-Ratings	Supervisory Ratings	Self-Ratings	Supervisory Ratings	Self-Ratings
Work quality	3.69	3.94	3.62	3.94	3.53	3.74	3.84	4.10
Work quantity	3.25	3.87	3.37	3.56	3.63	3.68	3.47	3.79
Judgment	3.25	3.81	3.31	3.87	3.31	4.05	3.52	3.79
Initiative	3.37	3.87	3.50	3.87	3.53	3.95	3.73	4.05
Teamwork	3.56	4.12	3.75	3.94	3.63	4.00	3.74	3.95
Dependability	3.44	3.94	3.56	3.87	3.63	4.10	3.79	4.16

[a] 16 men and 19 women supervisors; 35 men and 35 women subordinates.

was significant, we performed separate univariate analyses of variance tests for each performance dimension to determine which dimensions contributed to the significant multivariate result. A multivariate analysis of variance showed a significant main effect for source ($F_{6,61} = 4.13, p < .01$), indicating a general trend of greater leniency in self-ratings ($\bar{x} = 3.92$) than in supervisory ratings ($\bar{x} = 3.55$). Mean self-ratings were higher than mean supervisory ratings for the six performance dimensions for all combinations of supervisor's gender and subordinate's gender. Univariate analyses of variance (Table 2) showed a significant main effect for source for each of the performance dimensions.

The interaction between source and gender of subordinate was not significant (multivariate $F_{6,61} = .46$), indicating that subordinate's gender did not affect leniency of self-ratings relative to supervisory ratings. Because this multivariate result was not significant, we conducted no univariate analyses. As predicted in Hypothesis 1, men rated themselves higher than their supervisors rated them. Contrary to the predictions of Hypothesis 1, women did not rate themselves lower than their supervisors rated them. In fact, women tended to rate themselves higher than their supervisors rated them. The nonsignificant interaction for source × gender of subordinate also shows that subordinate's gender did not affect self-ratings, as predicted in Hypothesis 2. Moreover, men and women subordinates rated themselves similarly for all of the performance dimensions.

The interaction between source and supervisor's gender was not significant (multivariate $F_{6,61} = .23$). This indicates that supervisor's gender did not, as predicted in Hypothesis 3, affect leniency of self-ratings relative to supervisory ratings. Both men and women supervisors rated their subordinates lower than subordinates rated themselves. Further, the nonsignificant interaction between source and supervisor's gender indicates that there were no significant differences in the ratings made by men and women supervisors, contrary to predictions in Hypothesis 4.

An exploratory area of research was the potential three-way interaction for source × supervisor's gender × subordinate's gender. This interaction was

TABLE 2
Comparison of Mean Differences between Self- and Supervisory Ratings[a]

	Self-Ratings		Supervisory Ratings		
Performance Dimensions	**Means**	**s. d.**	**Means**	**s. d.**	***F***
Work quality	3.93	.77	3.67	.65	7.40*
Work quantity	3.67	.65	3.44	.79	7.98*
Judgment	3.89	.73	3.36	.82	19.19**
Initiative	3.94	.76	3.54	.76	15.24**
Teamwork	4.00	.74	3.67	.76	8.61*
Dependability	4.03	.76	3.61	.92	15.08**

[a] N = 70
*p<.01
**p<.001

not significant (multivariate $F_{6,61} = 1.26$), indicating that there were no differences in self- or supervisory ratings between same-gender and cross-gender supervisor/subordinate dyads.

The focus of the present study was the level of agreement between self- and supervisory ratings, which we tested by the effect called source. The main effects for gender of supervisor and gender of subordinate were not meaningful since they combined self- and supervisory ratings. For example, for the main effect for gender of supervisor, we combined ratings done by men supervisors with their subordinates' self-ratings to make one group and then compared these with a second group consisting of women supervisors' ratings combined with their subordinates' self-ratings. The same problem existed for the main effect gender of subordinate. In any event, these effects were not found to be significant.

DISCUSSION

Consistent with previous research (Thornton, 1980), the significant main effect for source indicates that subordinates tended to rate themselves higher than their supervisors rated them. The effect was consistent for both men and women subordinates regardless of supervisor's gender. This suggests that none of the variables investigated in this study influence leniency of subordinate ratings relative to supervisory ratings.

The lack of gender differences in both self- and supervisory ratings in the present study is not consistent with most previous research (Lenney, 1977; Maccoby & Jacklin, 1974; Nieva & Gutek, 1980). However, most previous research took place in laboratory settings whereas the present study had an industrial setting. Perhaps situational factors such as prior feedback that are present in applied settings but not in laboratories account for this discrepancy. The literature suggests two potential explanations of why the present study of working employees and supervisors might have produced ratings that were inconsistent with previous laboratory research: (1) gender differences in self-confidence, as expressed in performance ratings, occur in novel situations; and (2) stereotypes influence self- and supervisory ratings in situations for which information on performance is lacking.

Maccoby and Jacklin (1974) suggested that women are lower in self-confidence than men in many achievement situations. Their conclusion is based on the consistently lower performance expectancies and self-evaluations of both abilities and completed performance exhibited by women than by men. They thus assumed that lower self-ratings reflected lower self-confidence. However, Lenney (1977) suggested that gender differences in self-confidence are not likely to occur when individuals obtain clear information on their abilities to do specific tasks. A subsequent study by Lenney, Browning, and Mitchell (1980) concluded that women state lower expectancies than men when estimating their future performance on unfamiliar tasks, but do not do so for tasks on which they have received prior performance feedback. Thus, the effects of prior feedback on men's and women's self-confidence may explain the inconsistency of our findings on self-ratings with the findings of previous laboratory research.

The lack of gender differences we found may also be a result of women's low self-confidence on unfamiliar tasks. Prior research showing gender differences in ratings of others used undergraduates as subjects in simulated work settings (Hamner et al., 1974; London & Poplawski, 1976). The rating of others was a novel task for the undergraduates in those studies. Perhaps the higher ratings of others made by women compared to men reflected women's low self-confidence in their abilities to rate effectively in novel situations. In contrast, all of the supervisors participating in the present study had prior experience in performance rating. Both men and women supervisors were thus equally likely to feel confident in their abilities to rate others.

Gender roles stereotypes may have an important influence on performance ratings in some settings (Nieva & Gutek, 1980; Terborg & Ilgen, 1975). Terborg and Ilgen suggested that the effect of stereotypes is most potent in situations in which actual performance information is lacking and that stereotyping decreases as information is obtained. People use stereotypes to infer information about performance in the absence of better information. Lenney and her coauthors (1980) suggested that without external feedback, society's stereotype that women are less competent than men across a variety of situations is likely to influence self-evaluations. Therefore, women's low self-ratings in laboratory research may be due to their use of stereotypes in making ratings in the absence of performance feedback. In contrast, neither Snyder and Bruning (1979) nor the present study found gender differences in self-ratings; both these studies involved the behavior of men and women in actual work settings where feedback on performance was provided.

Performance feedback may not be the only variable that differentially influences men's and women's self-evaluations. Two field studies suggest variables that may interact with gender to affect self-ratings even when people have had prior performance feedback. First, status differences between supervisors and subordinates may produce low self-ratings among women. Waldman's (1979) study in the university setting of women secretaries' self-ratings compared with their faculty supervisors' ratings found that self-ratings were lower than supervisory ratings. In the present study, the first-level supervisors who participated had previously been assemblers, so we would not expect status differences.

Second, a job's gender-appropriateness may interact with gender to affect self-ratings. Deaux's (1979) study showing lower self-evaluations by women than men used first-level supervisors as subjects. Supervisory positions have been traditionally held by men. The assemblers participating in the present study were not in a stereotypically masculine field, and self-ratings of men and women were not different.

Use of stereotypes may also explain the lack of gender differences in supervisory ratings in the present study. Society's common view of women as nurturing, warm, and emotionally supportive (O'Leary, 1974) may influence women's ratings of others when they lack performance information.

Thus, women may rate others more leniently than men do in laboratory experiments or in selection decisions for which performance data are lacking. In contrast, actual work settings provide extensive performance information for evaluating current employees so that stereotypes should not influence supervisory ratings.

An important finding of the present study was the lack of a three-way interaction between supervisor's gender, subordinate's gender, and source of ratings. This is consistent with the results of other field studies that have found no interaction between gender of supervisor and gender of subordinate. Two studies showed no such interactions for supervisory ratings of subordinate's job performance (Mobley, 1982; Peters, O'Connor, Weekley, Pooyan, Frank, & Erenkrantz, 1984). A third (Wexley & Pulakos, 1982) showed no such interaction for ratings of supervisors by their subordinates or attitudinal congruence between supervisors and subordinates. Therefore, it appears that interactions between gender of supervisor and gender of subordinate may not affect performance ratings of actual employees.

A great deal of the preceding discussion focused on potential reasons for the nonsignificant effects of gender on self- and supervisory ratings in the present study. Another pertinent issue is the power of the statistical analyses. We conducted a power analysis (Cohen, 1977) to determine whether the present study had adequate power for detecting significant effects, should they exist.

Power analyses showed that for the main effect, source, and for the two- and three-way interactions between source and gender of supervisor and gender of subordinate, power was quite adequate for medium (.50) and large (.80) effect sizes. Power ranged from .79 to .995. If the size of an effect were medium or large, it is highly likely that the present analyses would have detected differences between men and women respondents, and between supervisors' ratings and self-ratings. However, if the effect size were small (.20), only the main effect for source of ratings had sufficient power to insure the likelihood of detecting group differences (power = .665). For the interactions, if the effect size were small, power was only .39 for source × gender of subordinate, .36 for source × gender of supervisor, and .19 for source × gender of subordinate × gender of supervisor.

In sum, the present study and other studies investigating interactions between supervisors' genders and subordinates' genders in real work settings have a potential limitation. Small effects may not be detectable owing to limitations of the populations being studied. Unfortunately, there are probably few work situations in which the number of women supervisors with men subordinates would allow the detection of small-sized effects. Laboratory studies may therefore be detecting small gender differences that are not discernable in applied research. Although small gender differences should certainly be documented, larger gender differences are of greater practical significance to organizations. The present research suggests that if gender does influence performance ratings, that effect is in all likelihood small.

One implication of the present study is that gender differences in self- and supervisory ratings may occur only in novel situations. With feedback on skill at performing a task, whether it be performance on the task itself or rating others' performance, gender differences in self-ratings and ratings of others may not occur. Thus, it is important that future research examine the effects of feedback on self- and supervisory ratings. Further, an investigation of whether self-confidence or stereotyping are related to gender differences in rating would be valuable.

Another area of future research the current study suggests is an examination of situational variables that may affect the relationship between self- and supervisory ratings. Personal characteristics such as gender may not have the direct relationship with ratings that many laboratory studies seem to suggest. Gender differences in ratings may operate only in some work situations—for example, when there are status differences, when a job is stereotypically masculine or feminine, or in certain occupations. The present study seems to suggest that situational variables such as prior experience or amount of previous feedback may interact with gender to affect the leniency of self-ratings compared to supervisory ratings.

CONCLUSIONS

This study clearly supports the value of testing laboratory findings in work settings. Since work environments are much more complex than laboratories it is important to determine whether conclusions based on lab studies are applicable in work settings. Of course, every work situation is unique, and results from research in one work setting may not apply to others. It is especially critical to evaluate the generalizability of studies of perceived performance since perceptions are likely to be influenced by aspects of the situation such as prior feedback, gender-typing of a job, or organizational type.

There are several important implications of this study for managers. First, managers need to be aware that employees tend to evaluate their own performances more highly than their supervisors evaluate them. Even when organizations do not use self-ratings as part of their appraisal process, employees will still evaluate their supervisors' appraisals relative to their own perceptions of their job performance. Discrepancies between self- and supervisory evaluations may well damage employees' self-esteem. Communication between managers and employees regarding perceptions each has of employee job performance should help remedy discrepancies. Second, managers need to evaluate gender differences in performance ratings for their particular setting. Until more research is done assessing the generalizability of the present findings across other organizations and jobs, no firm conclusions can be made regarding gender differences in performance ratings among employees. Finally, managers may be able to minimize gender differences in performance ratings by providing employees with accurate and realistic feedback about work-related performance, or by encouraging day-to-day interactions between employees and their supervisors so that supervisory evaluations are based on performance-related information rather than stereotypes.

REFERENCES

Baird, L. S. 1977. Self and superior ratings of performance: As related to self-esteem and satisfaction with supervision. ***Academy of Management Journal***, 20: 291–300.

Bartlett, C. J. 1959. The relationship between self-ratings and peer ratings on a leadership behavior scale. ***Personnel Psychology***, 12: 237–246.

Bassett, G. A., & Meyer, H. H. 1968. Performance appraisal based on self-review. ***Personnel Psychology***, 21: 421–430.

Bigoness, W. J. 1976. Effect of applicant's sex, race, and performance on employers' performance ratings: Some additional findings. ***Journal of Applied Psychology***, 61: 80–84.

Brief, A. P., Aldag, R. J., & VanSell, M. 1977. Moderators of the relationship between self and superior evaluations of job performance. ***Journal of Occupational Psychology***, 50: 129–134.

Cohen, J. 1977. ***Statistical power analysis for the behavioral sciences.*** New York: Academic Press.

Deaux, K. 1979. Self-evaluation of male and female managers. ***Journal of Sex Roles***, 5: 571–580.

Deaux, K., & Farris, E. 1977. Attributing causes for one's own performance: The effects of sex, norms, and outcome. ***Journal of Research in Personality***, 11: 59–72.

Dipboye, R. L., & Flanagan, M. F. 1979. Research settings in industrial and organizational psychology. ***American Psychologist***, 34: 141–150.

Elmore, P. B., & LaPointe, K. A. 1974. Effects of teacher sex and student sex on the evaluation of college instructors. ***Journal of Educational Psychology***, 66: 386–389.

Elmore, P. B., & LaPointe, K. A. 1975. Effects of teacher sex, student sex, and teacher warmth on the evaluation of college instructors. ***Journal of Educational Psychology***, 67: 368–374.

Fidell, L. S. 1970. Empirical verification of sex discrimination in hiring practices in psychology. ***American Psychologist***, 25: 1094–1098.

Frank, F. D., & Drucker, J. 1977. The influence of evaluatee's sex on evaluation of a response of a managerial selection instrument. ***Sex Roles***, 3: 59–64.

Hall, F. S., & Hall, D. T. 1976. Effects of job incumbents' race and sex on evaluations of managerial performance. ***Academy of Management Journal***, 19: 476–481.

Hamner, W. C., Kim, J. S., Baird, L., & Bigoness, W J. 1974. Race and sex as determinants of ratings by potential employers in a simulated work sampling task. ***Journal of Applied Psychology***, 59: 705–711.

Heneman, H. G., III. 1974. Comparisons of self- and superior-ratings of managerial performance. ***Journal of Applied Psychology***, 59: 638–642.

Holzbach, R. L. 1978. Rater bias in performance ratings: Superior, self-, and peer ratings. ***Journal of Applied Psychology***, 63: 579–588.

Jacobson, M. B., & Effertz, J. 1974. Sex roles and leadership: Perceptions of the leaders and the led. ***Organizational Behavior and Human Performance***, 12: 383–396.

James, L. R. 1973. Criterion models and construct validity for criteria. ***Psychological Bulletin***, 80: 75–83.

Kirchner, W. K. 1966. Relationships between supervisory and subordinate ratings for technical personnel. ***Journal of Industrial Psychology***, 3: 57–60.

Kirk, R. E. 1982. ***Experimental design: Procedures for the behavioral sciences*** (2nd ed.). Belmont, Calif.: Brooks/Cole.

Klimoski, R. J., & London, M. 1974. Role of the rater in performance appraisal. ***Journal of Applied Psychology***, 59: 445–451.

Landy, F. J., & Farr, J. L. 1980. Performance rating. ***Psychological Bulletin***, 87: 72–107.

Lawler, E. E., III. 1967. The multitrait-multirater approach to measuring managerial job performance. ***Journal of Applied Psychology***, 51: 369–381.

Lee, D. M., & Alvares, K. M. 1977. Effects of sex on descriptions and evaluations of supervisory behavior in a simulated industrial setting. ***Journal of Applied Psychology***, 62: 405–410.

Lenney, E. 1977. Women's self-confidence in achievement settings. ***Psychological Bulletin***, 84: 1–13.

Lenney, E., Browning, C., & Mitchell, L. 1980. What you know can hurt you: The effects of performance criteria ambiguity on sex differences in self-confidence. ***Journal of Personality***, 48: 306–322.

Levine, E. L., Flory, A., & Ash, R. A. 1977. Self-assessment in personnel selection. ***Journal of Applied Psychology***, 62: 428–435.

London, M., & Poplawski, J. R. 1976. Effects of information on stereotype development in performance appraisal and interview contexts. ***Journal of Applied Psychology***, 61: 199–205.

Maccoby, E. E., & Jacklin, C. N. 1974. ***The psychology of sex differences***. Stanford, Calif.: Stanford University Press.

Meyer, H. H. 1977. The annual performance review discussion—Making it constructive. ***Personnel Journal***, 56: 508–511.

Mobley, W. H. 1982. Supervisor and employee race and sex effects on performance appraisals: A field study of adverse impact and generalizability. ***Academy of Management Journal***, 25: 598–606.

Nieva, V. F., & Gutek, B. A. 1980. Sex effects on evaluation. ***Academy of Management Review***, 5: 267–276.

O'Leary, V. E. 1974. Some attitudinal barriers to occupational aspirations in women. ***Psychological Bulletin***, 81: 809–826.

Parker, J. W., Taylor, E. K., Barrett, R. S., & Martens, L. 1959. Rating scale content: III. Relationship between supervisory and self-ratings. ***Personnel Psychology***, 12: 49–63.

Peters, L. H., O'Connor, E. J., Weekley, J., Pooyan, A., Frank, B., & Erenkrantz, B. 1984. Sex bias and managerial evaluations: A replication and extension. ***Journal of Applied Psychology***, 69: 349–352.

Pheterson, G. T., Kiesler, S. B., & Goldberg, P. A. 1971. Evaluation of the performance of women as a function of their sex, achievement, and personal history. ***Journal of Personality and Social Psychology***, 19: 114–118.

Primoff, E. S. 1980. The use of self-assessments in examining. ***Personnel Psychology***, 33: 283–290.

Rosen, B. R., & Jerdee, T. H. 1974. Influence of sex role stereotypes on personnel decisions. ***Journal of Applied Psychology***, 59: 9–14.

Schmitt, N., & Lappin, M. 1980. Race and sex as determinants of the mean and variance of performance ratings. ***Journal of Applied Psychology***, 65: 428–435.

Snyder, R. A., & Bruning, N. S. 1979. Sex differences in perceived competence: An across organizations study. ***Administration in Social Work***, 3: 349–358.

Taynor, J., & Deaux, K. 1973. When women are more deserving than men: Equity, attribution, and perceived sex differences. ***Journal of Personality and Social Psychology***, 28: 360–367.

Terborg, J. R. 1977. Women in management: A research review. ***Journal of Applied Psychology***, 62: 647–664.

Terborg, J. R., & Ilgen, D. R. 1975. A theoretical approach to sex discrimination in traditionally masculine occupations. ***Organizational Behavior and Human Performance***, 13: 352–376.

Thornton, G. C., III. 1968. The relationship between supervisor- and self-appraisals of executive performance. ***Personnel Psychology***, 21: 441–445.

Thornton, G. C., III. 1980. Psychometric properties of self-appraisals of job performance. ***Personnel Psychology***, 33: 263–271.

Waldman, D. A. 1979. ***A comparison of self- and supervisor-appraisals moderated by level of anonymity and perceptions of organizational climate***. Unpublished master's thesis, Colorado State University, Fort Collins.

Waldman, D. A., & Thornton, G. C., III. 1979. A comparison of supervisors' self-appraisals and their administrators' appraisals. ***Medical Group Management***, 22, 24, 58.

Wexley, K. N., & Pulakos, E. D. 1982. Sex effects on performance ratings in manager-subordinate dyads: A field study. ***Journal of Applied Psychology***, 67: 433–439.

Lynn McFarlane Shore earned her Ph.D. degree in industrial/organizational psychology at Colorado State University, Fort Collins, Colorado; she is Visiting Assistant Professor of Management in the Department of Management and Labor Relations at Cleveland State University, Cleveland, Ohio. Her current research interests include work commitment, performance appraisal, and management development.

George C. Thornton III earned his Ph.D. degree in industrial psychology at Purdue University; he is Professor of Industrial/Organizational Psychology at Colorado State University, Fort Collins, Colorado. His current research interests include performance appraisal, managerial assessment centers, and evaluation of management training programs.

RESEARCH NOTES

© *Academy of Management Journal*
1986, Vol. 29, No. 1, 130–139.

USING THE LENS MODEL TO MAGNIFY RATERS' CONSISTENCY, MATCHING, AND SHARED BIAS

NEAL SCHMITT
Michigan State University
RAYMOND A. NOE
University of Minnesota
RAND GOTTSCHALK
Michigan State University

Performance ratings have received increased attention in the last decade partly because of their importance for many personnel functions, but also because the judgments required are interesting cognitive phenomena (Ilgen & Feldman, 1983). Differences among the ratings given by supervisors, peers, and subordinates are commonly noted (Kavanagh, McKinney, & Wolins, 1971). These differences could be due to a variety of reasons, including (1) differing opportunities to observe behavior and (2) different types of raters using different cognitive schemata to evaluate performance. Supervisors, peers, and subordinates are likely to have different frameworks for integrating information concerning performance, which results in variance in dimensional and overall performance ratings within and across rating groups.

Policy-capturing techniques have been used to describe how raters combine information on individual performance dimensions to arrive at overall ratings of performance (Hobson, Mendel, & Gibson, 1981; Zedeck & Kafry, 1977). Capturing the policies of various groups of raters also allows comparisons concerning how different raters or rating groups use each type of information in making performance ratings. Understanding how people use certain types of information can pinpoint more precisely what kinds of factors to emphasize when training raters.

Much of the policy-capturing literature originated with a model of perceptual process presented by Brunswik (1952). Brunswik was interested in how people's perceptions of their world coincided, or failed to coincide, with objective reality. He presented his ideas pictorially as a lens model. Policy-capturing research generated by this lens model has ignored various aspects of the model that may be of potential value in examining interrater agreement among raters or rater groups. Specifically, we argue that lack of agreement may occur for two reasons. The first is that individuals do not

---

The financial assistance of the Spencer Foundation, the National Association of Secondary School Principals, and the Rockefeller Family Fund in the preparation of this paper is gratefully acknowledged. Critical comments by Steve Kozlowski and Daniel Ilgen on an earlier version of this manuscript were also helpful.

apply any policy consistently. Consistency is the problem usually investigated by policy-capturing research. Second, lack of agreement may occur because individuals fail to apply similar policies. We further argue that interrater agreement may be too high when it exceeds values predicted by the lens model because of shared rater biases. Finally, scholars have criticized policy-capturing research for using hypothetical performance profiles rather than data collected in actual performance-rating sessions (Hobson & Gibson, 1983).

In our Methods section we describe a methodology based on the lens model that allows analysis of (1) the degree to which raters adopt similar strategies of combining information, or matching; and (2) the degree to which their agreement is a function of job-relevant considerations or a function of shared biases. We then provide an empirical example and analyze it using these lens model indices.

METHODS

Lens Model Conceptualization of Interrater Agreement

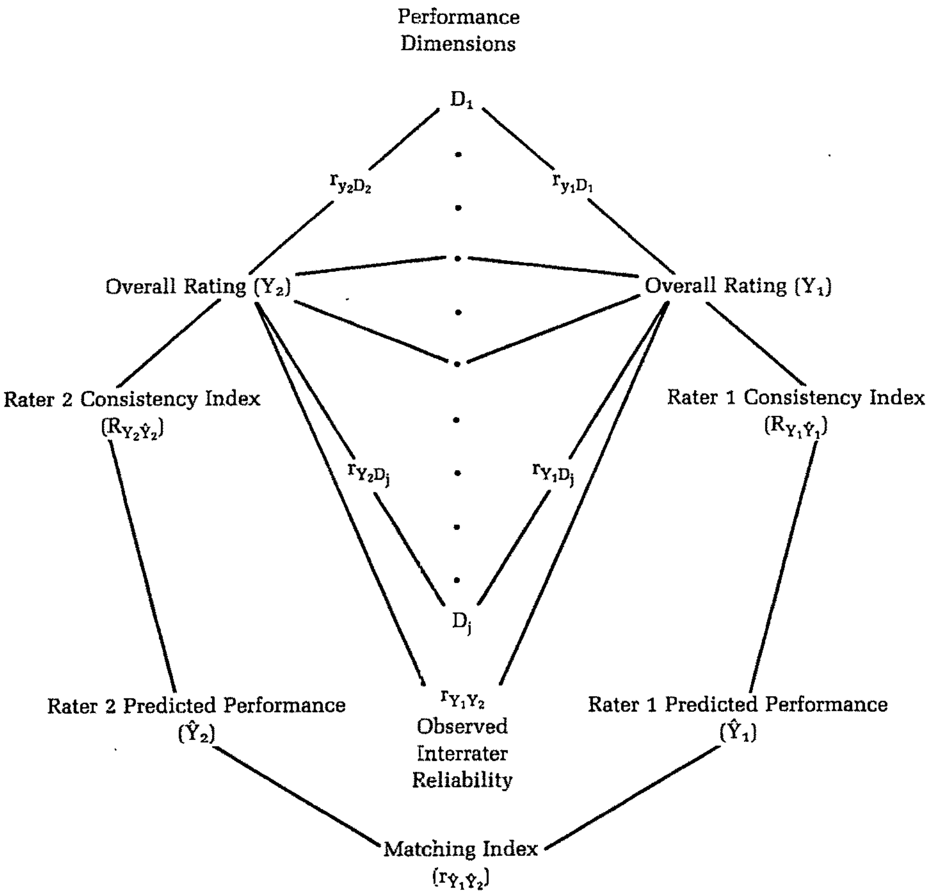
Figure 1 presents a pictorial representation of our adaptation of the lens model for two raters. Multiple raters or rater groups would simply involve more indices of the same type as appear on either side of the lens model. One major assumption underlying the lens model is that raters' judgments of ratees' job performance reflect the same individual performance dimensions relevant to the overall performance of a job. An additional assumption is that a linear model represents all judgments adequately. Individual performance is rated on a number of dimensions ($D_1 \ldots D_N$). In our example, the individual values of $r_{y_i D_j}$ represents correlations between performance ratings on major job performance dimensions and an overall rating of job performance, where y_i corresponds to the overall performance rating and D_j equals the rating given to the jth dimension. The multiple correlation resulting from the regression of a rater's overall performance ratings on measures of individual performance dimensions represents the degree to which information on these individual dimensions contributes to assessments of overall job performance. This provides an index of consistency. Using the regression equations to predict values of overall performance ($\hat{Y}_1$ and $\hat{Y}_2$) and then correlating these two sets of predicted values yields a matching index ($r_{\hat{Y}_1\hat{Y}_2}$). This index, largely ignored in policy-capturing research focusing on performance ratings, represents the degree to which decision makers use similar strategies when they are combining information.

For the purposes of this study, we adapted Slovic and Lichtenstein's (1971: 657) lens model equation. Our formula for the observed correlation between ratings is as follows:

$$r_{Y_1Y_2} = r_{\hat{Y}_1\hat{Y}_2} R_{Y_1\hat{Y}_1} R_{Y_2\hat{Y}_2} \quad . \tag{1}$$

Equation 1 states that the correlation between two raters' judgments of a ratee's overall performance is a multiplicative function of (1) the degree to which the two raters use similar strategies of combining information ($r_{\hat{Y}_1\hat{Y}_2}$)

FIGURE 1
Lens Model Representation of Interrater Reliability for Two Raters


Performance Dimensions
D_1
$r_{y_2D_2}$
$r_{y_1D_1}$
Overall Rating (Y_2)
Overall Rating (Y_1)
Rater 2 Consistency Index ($R_{Y_2\hat{Y}_2}$)
Rater 1 Consistency Index ($R_{Y_1\hat{Y}_1}$)
$r_{Y_2D_j}$
$r_{Y_1D_j}$
D_j
Rater 2 Predicted Performance ($\hat{Y}_2$)
$r_{Y_1Y_2}$ Observed Interrater Reliability
Rater 1 Predicted Performance ($\hat{Y}_1$)
Matching Index ($r_{\hat{Y}_1\hat{Y}_2}$)

and (2) the degree to which raters use available information consistently in coming to overall evaluations of performance ($R_{Y_1\hat{Y}_1}$ and $R_{Y_2\hat{Y}_2}$). Analyzing interrater reliability in this fashion provides information concerning several interesting practical questions. First, are decision makers using similar information in varying situations and for different purposes? Second, to what extent are decision makers using this information in coming to overall judgments of performance? Third, what explains differences in the decisions of different sets of decision makers—lack of consistent information use or failure to adopt similar strategies?

The possibility exists that observed interrater reliabilities will exceed those predicted in Equation 1. This could be due to curvilinear use of information when decision makers make overall judgments. Alternatively, it could represent some linear relationship among the residuals caused by the availability of similar information to various decision makers rating job performance. This would be information not correlated with the dimensions that are measured. Assuming that a relatively thorough job analysis produces an adequate criterion and a reasonably complete set of performance dimensions, we could most likely attribute such a linear relationship to bias. It is important to note that bias can occur in the individual dimensional ratings as well. Two raters, or two groups of raters, may make both overall judgments and dimension judgments with similar biases. Our analysis does not deal with that problem, which requires some type of independent dimensional true score (Borman, 1978).

One possible source of such dual contamination or bias may be common stereotypes associated with ratees' demographics. In our example, we estimated this contamination using the following equation:

$$r_{DRINC} = R_{DRES_{12}} [(1\text{-}R^2{}_{Y_1\hat{Y}_1}) (1\text{-}R^2{}_{Y_2\hat{Y}_2})]^{1/2} \quad , \tag{2}$$

where r_{DRINC} equals the incremental intercorrelation that occurs through the demographic correlates of the residuals and $R_{DRES_{12}}$ is the multiple correlation between the demographics and the two sets of residuals involved. The residual-demographic correlations result in increases in interrater correlations equal to the product of the multiple correlation between the residuals and the demographic variables and the residuals themselves.

Next, we present an example to demonstrate how policy-capturing techniques can be used with actual performance ratings to illuminate three issues. The first is consistency, how raters combine information on individual performance dimensions to make overall performance ratings. The second we call matching, the extent to which different rating groups use similar strategies in making judgments about levels of overall performance. The third is the degree to which common biases may inflate agreement on such overall judgments. In particular, we illustrate how we used the Brunswik (1952) lens model to analyze three different groups of raters' judgments of the overall performance of a group of school administrators.

Participants

Performance ratings were collected for 153 school administrators in five different school districts throughout the country. These ratings were used as criteria, for research purposes only, in the validation of a recently developed assessment center designed to select school administrators. These administrators included 91 men and 62 women, 129 of whom were white and 24 of whom were nonwhite.

Measures

Performance ratings were collected on 15 performance dimensions using behaviorally anchored rating scales. Raters were also asked to make overall performance ratings. Ratings of administrators' performance were completed by their immediate supervisors, two teachers, and the administrators themselves. Because of their high intercorrelation, the two teachers' ratings were averaged to provide a composite teacher rating. Table 1 presents short descriptions of the 15 performance dimensions. The full scales are available upon request from the senior author.

Data Analyses

The first step in the data analysis involved the regression of the overall performance rating on the performance dimensions for each of the rater groups. The other correlational indices ($r_{Y_1Y_2}$ and $r_{\hat{Y}_1\hat{Y}_2}$) suggested by Figure 1 were computed as well as the correlations between the residuals of the regression equations and candidates' races, genders, and positions.

RESULTS

Intercorrelations among ratings on the 15 performance dimensions were relatively high. The intercorrelations among self-rating dimensions ranged from .06 to .64 and averaged .35. The range of intercorrelations for the supervisor ratings was .45 to .79 with an average of .65. The average intercorrelation among teacher ratings was .69 and the range was .51 to .87. We do not present full intercorrelation matrices because of space considerations, but they are available from the senior author.

TABLE 1
Results of Regression Analysis of Overall Job Performance on Performance Dimensions for Ratings by Supervisor, Self, and Teachers

	Beta Weights for Prediction of Overall Judgments by:		
Performance Dimensions	**Self** (*n*=123)	**Supervisor** (*n*=86)	**Teacher** (*n*=79)
Monitoring curriculum objectives	.204*	.114	.139
Monitoring student progress	.193*	.013	.048
Supervision of student activities	.103	.148	.057
Participation in student activities	.093	−.026	−.039
Direction of support services	−.100	.181*	.104
Directing behavior of students	.004	.224*	.163
Staff evaluation	.108	−.052	.178*
Developmental activities	−.033	.111	.029
Community relations	.135	−.066	.003
Interpersonal effectiveness	−.023	.046	.007
Relationship with parents	.117	.051	.162*
Coordination with district	.305*	.011	.009
Fiscal management	−.194*	−.086	.035
Maintenance of school plant	.060	.063	.064
Structure communication	.120	.317*	.170
Multiple R	.79*	.88*	.93*
Multiple R^2	.62	.77	.86
Adjusted R^2	.57	.72	.82

*$p<.05$

Table 1 presents the results of the regression analysis for the three rater groups. All three sets of overall performance ratings were highly predictable on the basis of ratings of individual performance dimensions. The three vectors of regression weights were compared using the Chow (1960) test. These tests indicated that the regression coefficients associated with the supervisors' judgments were significantly different from those of the teachers' judgments ($F = 3.11$, $df = 16{,}133$, $p<.01$). Similarly, the regression equations for self- and supervisors' judgments were significantly different ($F = 3.30$, $df = 16{,}177$, $p<.01$) as were the equations for self- and teachers' judgments ($F = 6.99$, $df = 16{,}170$, $p<.01$).

Examination of the standardized regression coefficients from Table 1 indicates that various dimensions best predicted overall ratings by different groups of raters. The ratees themselves seemed to most heavily weight efforts directed at curriculum, student progress, fiscal matters, and district-school coordination problems. Supervisors'overall judgments were most related to judgments of ratees' efforts to direct support services and student behavior, and to structure communication. Teachers' ratings of the overall performance of their principals was most heavily influenced by the principals' efforts in staff evaluation and relationships with parents. An examination of the dimension descriptions and those dimensions given the most weight by different rater groups reveals that raters are weighting those dimensions that would seem most salient to them. For example, it seems reasonable that teachers would be concerned about staff evaluation and harmonious relationships with parents. However, we make all such statements cautiously, given the high degree of multicollinearity in the ratings.

In Table 2, we present the remainder of the correlations suggested by Figure 1. The observed correlations among overall ratings provided by different rater groups are low but about what would seem likely from the research literature. As Table 1 reported, the multiple correlations relating overall ratings to dimension ratings are uniformly high. However, the matching indices are low, which indicates that individuals in these different rater groups used different weights in combining information to come to judgments concerning overall performance of principals. Finally, predicted intergroup correlations using Equation 1 are lower than the observed correlations.

As suggested earlier, this discrepancy could occur because different sets of judges use information in a similar nonlinear fashion or because different sets of judges use information that the various performance rating dimensions do not include. To uncover a possible source of such dual contamination, we correlated five demographic variables with the residuals of those ratings. The demographic variables used in this analysis included the races and genders of candidates and three pieces of information concerning their current employment status: (1) the school level at which they were currently working, (2) whether they had been newly promoted from teaching, and (3) whether they had been assistant principals. The last two items were coded dichotomously.

TABLE 2
Lens Model Correlations

Rater Pairs	Consistency Index ($R_{Y_1\hat{Y}_1}$ and $R_{Y_2\hat{Y}_2}$)	Observed Intercorrelation ($r_{Y_1Y_2}$)	Matching Index ($r_{\hat{Y}_1\hat{Y}_2}$)	Predicted Overall Performance Intercorrelation (Equation 1)	Intercorrelation of Regression Residual and Demographic Variables (Equation 2)	Residual/Unexplained Intercorrelation[a]
Supervisor-self	.79 (self)	.31	.21	.19	.11	.01
Supervisor-teacher	.89 (supervisor)	.22	.20	.15	.07	.00
Self-teacher	.93 (teacher)	.41	.43	.36	.05	.00

[a] Computed by subtracting the predicted overall performance intercorrelation plus the intercorrelation of regression residuals and demographic variables from the observed intercorrelation.

The cross-rater correlates of residuals suggested by Equation 2 provide useful information about possible rater bias that is not relevant to measured job performance dimensions. The column headed "Equation 2" in Table 2 presents incremental intercorrelations among rater groups resulting from the cross products of the multiple correlations between demographic variables and the residuals. We used multiple correlations because of the demographics' lack of independence. As can be seen, substantial additional correlation represents an inflation of interrater correlation because both rating groups share similar biases.

SUMMARY AND CONCLUSIONS

In this paper, we have presented an interpretation in terms of the lens model of the relationship among ratings supplied by different groups of raters. We used the model (1) to determine the degree to which overall performance ratings were predictable from information available to the raters; (2) to examine similarities in the weighting systems used by different groups of raters; and (3) to examine potential inflation of interrater reliability through bias.

The analysis provides information that is useful for increasing the effectiveness of any performance appraisal system. Analysis of equations relating overall performance ratings to dimensional ratings provides two types of information. First, it reveals what rating dimensions most strongly influence overall performance ratings. Second, it shows the extent to which raters' overall evaluations of performance are based on dimensions other than those involved in appraisal forms. Because an overall performance rating is often administratively the most important piece of information in salary and promotion decisions, it is important that it be unbiased. Two conditions identified by use of the lens model may indicate a need for training of raters. If regression of the overall performance rating on dimensional ratings results in a small multiple R, it may indicate that raters are making overall ratings on the basis of other factors, such as gender or race, or that they are inconsistently using the dimensional information. Examination of matching indices as well as standardized regression weights allows for a comparison of different rater policies. Application of a different policy is psychologically different from inconsistent application of a policy and would dictate a different rater training strategy.

Besides identifying possible needs for rater training, the lens model analysis provides an objective statistical description of rater behavior. Sharing information with raters concerning (1) the extent to which their ratings are similar to other organizational members' ratings and (2) the importance of each performance dimension in determining their overall ratings can help them evaluate the extent to which their actual rating behavior reflects their intended behavior. Raters may modify their individual decision making processes to fit organizational desires if they receive objective analyses of how their rating behavior compares to that of others in their rating group and to organizational policy. However, this hypothesis requires empirical verification.

Examination of the residual correlations can explain bias. Demonstrating that the residuals of equations relating individual performance dimensions to overall appraisals of job performance are correlated through common demographic correlates could be used to explain or correct inflated interrater correlations. In this context, we should emphasize that finding common demographic correlates of the residuals would not preclude corresponding differences on both dimension ratings and overall ratings.

The model outlined in this study could be applied to other problems in the human resource area. If performance dimensions parallel those dimensions on which selection decisions are made, as they should in job-relevant systems, organizations could compare decisions concerning selection and performance appraisals the way we compared rater groups. Analysis of the correlates of the residuals of the selection and performance appraisal equations could reveal the source of predictor-related criterion bias. This problem has received little attention, though it has obvious implications for the fairness of selection decisions. Analysis of job evaluation decisions, job analysis data, and assessment center judgments are other potential areas of application.

REFERENCES

Borman, W. C. 1978. Exploring upper limits of reliability and validity in job performance ratings. ***Journal of Applied Psychology***, 63: 135–144.

Brunswik, E. 1952. ***The conceptual framework of psychology***. Chicago: University of Chicago Press.

Chow, G. C. 1960. Tests of equality between subsets of coefficients in two linear regressions. ***Econometrica***, 28: 591–605.

Hobson, C. J., & Gibson, F. W. 1983. Policy capturing as an approach to understanding and improving performance appraisal: A review of the literature. ***Academy of Management Review***, 8: 640–649.

Hobson, C. J., Mendel, R. M., & Gibson, F. W. 1981. Clarifying performance appraisal criteria. ***Organizational Behavior and Human Performance***, 28: 164–188.

Ilgen, D. R., & Feldman, J. M. 1983. Performance appraisal: A process approach. In B. M. Staw and L. L. Cummings (Eds.), ***Research in Organizational Behavior***, Vol. 5: 141–198. Greenwich, Conn.: JAI Press.

Kavanagh, M. J., McKinney, A. C., & Wolins, L. 1971. Issues in managerial performance: Multitrait-multimethod analyses of ratings. ***Psychological Bulletin***, 75: 34–49.

Slovic, P., & Lichtenstein, S. 1971. Comparison of Bayesian and regression approaches to the study of information processing in judgment. ***Organizational Behavior and Human Performance***, 6: 649–744.

Zedeck, S., & Kafry, D. 1977. Capturing rater policies for processing evaluation data. ***Organizational Behavior and Human Performance***, 18: 269–294.

Neal Schmitt earned his Ph.D degree at Purdue University. Currently, he is Professor of Psychology and Management at Michigan State University. His present research interests include various issues in personnel selection, test validation, and decision making.

Raymond A. Noe received his Ph.D. degree in industrial/organizational psychology from Michigan State University. He is currently an assistant professor of labor and industrial relations in the Institute of Labor and Industrial Relations at the University of Minnesota. His teaching, research, and consulting interests center in the areas of staffing, training, and development.

Rand Gottschalk is a Ph.D. student in industrial organizational psychology at Michigan State University. His research interests include meta-analysis and organizational implementation of innovations.

© *Academy of Management Journal*
1986, Vol. 29, No. 1, 139–149.

EFFECTS OF TASK DIFFICULTY AND INTERUNIT INTERDEPENDENCE ON INFORMATION PROCESSING SYSTEMS

JACK K. ITO
University of Regina
RICHARD B. PETERSON
University of Washington

The roles of task uncertainty within subunits and interunit interdependence in explaining variations in organizational designs have been subject to much theorizing (Fry, 1982; Gerwin, 1981; McCann & Galbraith, 1981). Previous research has found task uncertainty in subunits to be related (1) to personal means of intraunit coordination (Hrebiniak, 1974; Mohr, 1971; Pennings, 1975a; Tushman, 1979a; Van de Ven, Delbecq, & Koenig, 1976); (2) to boundary-spanning activity (Leifer & Huber, 1977; Tushman, 1977); and (3) to such facets of intraunit structuring as participation in decision making (Mohr, 1971) and autonomy in task performance (Schoonhoven, 1981). Perrow (1967) identified two components of task uncertainty—number of exceptions and analyzability—that are conceptually distinct but empirically related (Withey, Daft, & Cooper, 1983). The present study focuses on the analyzability component of task uncertainty, which we labeled task difficulty. Task difficulty refers to knowledge of cause-effect relationships (Van de Ven & Delbecq, 1974). Part of this study relates task difficulty to boundary-spanning activity, participation, and autonomy.

The authors wish to thank Louis Fry, Kalburgi Srinivas, and two anonymous reviewers for their helpful comments on earlier drafts of this paper.

Little empirical work on the relationship between interdependence and coordination mechanisms exists (Cheng, 1983). Only Tushman (1977, 1978) studied linkage between interunit interdependence and boundary-spanning activity; these studies concerned relationships between laboratory units and other units in a high technology firm. The present study also relates interdependence between units to interunit boundary-spanning activity in the context of an administrative system; it therefore complements Tushman's research.

No study has specifically addressed the relationship between boundary-spanning activity, one mechanism that coordinates relations between units, and elements of intraunit structuring. Research addressing these linkages has either combined forms of boundary-spanning activity or has used global measures of interaction (Aiken & Hage, 1968; Bacharach & Aiken, 1977; Hrebiniak, 1974; Leifer & Huber, 1977; Mohr, 1971). Finally, this study relates boundary-spanning activity to participation in decision making and to autonomy.

THEORETICAL ANALYSIS AND HYPOTHESES

Information processing, defined as "the gathering, interpreting, and synthesis of information in the context of organizational decision-making" (Tushman & Nadler, 1978: 614), has served to explain variations in organization designs. The capacity to process information is a function of a unit's design that includes (1) mechanisms linking units and (2) intraunit structuring. Task difficulty and interunit interdependence are two factors that determine requirements for information processing capacity. Because requirements and capacity should match, these factors are linked to the designs of work units (Galbraith, 1973, 1977; Tushman & Nadler, 1978).

Boundary Spanning Activities

Organizational subunits manage their interunit interdependencies through mechanisms that vary in information-processing capacity (McCann & Ferry, 1979; Thompson, 1967). Programming, although the least costly mechanism, can only deal with predictable events in which cause-effect relations are clear (Connolly, 1977; Galbraith, 1972). Supervisors, acting as boundary spanners, may deal with interunit decisions not falling within procedures. An alternative mechanism, direct lateral contact among subordinates, may provide less control and be more costly than supervisory boundary spanning (Galbraith, 1972). Coding or linguistic barriers (Tushman, 1979a) may also hinder contact among subordinates. However, if information-processing requirements exceed supervisors' capacities, organizations use supplementary mechanisms (Thompson, 1967; Tushman & Nadler, 1978; Van de Ven et al., 1976). These mechanisms include boundary spanning by unit members through gatekeepers, in instances in which differentiation is high (Tushman, 1977; Tushman & Scanlan, 1981), and boundary spanning through direct contact where such barriers are low (Leifer & Huber, 1977).

Such transactions may range from one-on-one meetings to joint problem-solving teams (Van de Ven et al., 1976) and may be assisted by third parties (Lawrence & Lorsch, 1967).

McCann and Ferry stated, "Interdependence exists when actions taken by one referent system affects [*sic*] the actions or outcomes of another referent system" (1979: 113). In this paper, our concern was the degree of policy and procedural interdependence between focal units and other units as indicated by changes that other units induced in focal units.[1] The greater the interdependence, the greater would be the need to gather, analyze, and distribute information between subunits. Boundary-spanning activity by unit members, supplementing the limited capacity of rules and hierarchy (Tushman, 1977, 1978), may meet this need for information processing. Therefore, we expect that:

> *Hypothesis 1: The greater the degree of interunit interdependence, the greater the amount of boundary-spanning activity by unit members.*

It is a central theme in previous research that the greater the task difficulty, the greater the requirement for information-processing capacity (Duncan, 1973; Galbraith, 1977; Tushman & Nadler, 1978). Lateral communications that supplement programming and hierarchy in providing advice and information to unit members in task performance may meet such requirements (Burns & Stalker, 1961). These networks would provide (1) reduced equivocality of information (Daft & Macintosh, 1981); (2) conceptual frameworks (Connolly, 1977); (3) social validation of decisions (Connolly, 1977; Litwak, 1961); and (4) coordination (March & Simon, 1958; Thompson, 1967).

When interunit interdependence is low, these networks may be restricted to intraunit transactions, with supervisors managing interfaces between units. However, with high levels of interdependence, expanding networks to include members from other units may facilitate needed extraunit information exchange (Galbraith, 1972). Therefore, since the level of interunit interdependence for most of the work units studied was moderate or high, we propose that:

> *Hypothesis 2: The greater the degree of task difficulty, the greater the amount of boundary-spanning activity by unit members.*

Intraunit Structuring

Previous research has developed and tested linkages between task difficulty and participation and has found varying levels of support (Hage & Aiken, 1969; Hrebiniak, 1974; Grimes, Klein, & Schull, 1972; Mohr, 1971). Briefly, participation by subordinates enhances supervisory capacity to generate and evaluate alternatives (Duncan, 1973). The more difficult the tasks, the greater the vesting of cause-effect knowledge in subordinates, hence the greater their potential contributions to decision making. Therefore, we hypothesize that:

---

[1] For other frameworks, see McCann and Ferry (1979); Pennings (1975b); Thompson (1967); and Van de Ven et al. (1976).

> *Hypothesis 3: The greater the degree of task difficulty, the greater the level of participation in decision making by unit members.*

The relationship between task difficulty and discretion at the point of task execution enjoys considerable theoretical and empirical support (Bell, 1967; Burns & Stalker, 1961; Grimes et al., 1972; Hrebiniak, 1974; Schoonhoven, 1981; Thompson & Bates, 1957). When task difficulty is low, subordinates may refer exceptions upward; but as cause-effect relationships become less tangible, units may expand information processing capacity by increasing discretion at lower levels. Granting authority to subordinates to make decisions within prescribed boundaries serves at least three functions: (1) it prevents overloading supervisors (Katz & Kahn, 1978); (2) it improves timeliness of responses; and (3) it places decisions at levels at which there may be greater expertise about particular issues than supervisors have. Control of subordinates may be exercised through professionalization, the inculcation of norms and values, or the specification of goals (Galbraith, 1973). Therefore, we expect that:

> *Hypothesis 4: The greater the degree of task difficulty, the greater the degree of autonomy of unit members.*

Boundary-spanning activities imply both joint decision making with members of other units and needs to seek out and consolidate extraunit information in performing daily tasks. Thus levels of boundary-spanning activity are likely to be associated with the extent of subordinates' knowledge regarding other units. Where extraunit knowledge is dispersed, rather than concentrated among supervisors, there are potential subordinate roles in making decisions having extraunit considerations in at least two situations. First, supervisors would be partly dependent on subordinates for information and advice when making decisions with extraunit considerations. Although they differ in the form of boundary-spanning activity investigated, several studies (Aiken & Hage, 1968; Hrebiniak, 1974; Leifer & Huber, 1977; Mohr, 1971; Tushman, 1979a) have supported its linkage to participation. Second, such joint decision making and subordinate expertise in extraunit matters would be consistent with limited supervision in day-to-day decisions with extraunit considerations. The work of Aiken and Hage (1968), Bacharach and Aiken (1977), and Leifer and Huber (1977) supported this linkage. Therefore, we expect that:

> *Hypothesis 5: The greater the degree of boundary-spanning activity, the greater the level of unit members' participation in decision making.*
>
> *Hypothesis 6: The greater the degree of boundary-spanning activity, the greater the degree of unit members' autonomy.*

METHODS

Research Setting and Sample

The senior author carried out this study in the field and head office units of the Department of Employment Security of a western state as an unpaid researcher. Van de Ven and Delbecq (1974) described this agency's functions and structure. Following Van de Ven and colleagues (1976) and Tushman (1977, 1978), we used work groups as the units of analysis. First-line supervisors with at least one year as unit heads who had full-time subordinates with sufficient experience to be familiar with their positions answered questions on the subordinates' jobs. The study included all units meeting the criterion for supervisory experience and containing at least two experienced subordinates. Each potential respondent received an individually addressed questionnaire through the internal mail system. Respondents mailed completed questionnaires directly to the senior author. The data analysis includes only units in which responses came from both the supervisor and from at least two subordinates representing at least one-third of those qualified to respond. Questionnaires went out to 612 potential respondents in 131 work units; 381 questionnaires (62%) were returned. Seventy-five units met the criteria for data analysis; these units included 360 potential respondents of whom 283 responded for a 79 percent response rate.

Variables

Brief descriptions of each of the variables follow.[2] All items were measured on 7-point Likert scales. Factor analysis with varimax rotation was used to group items into variables. The items selected loaded at > .40 on one factor. Table 2 gives means, standard deviations, and reliabilities.

The four task difficulty items followed Perrow's (1967) concept of analyzability (e.g., the extent to which individuals use standard methods). The two interunit interdependence items concern change in policies and procedures induced by other subunits, a concept that applies to both B→ A and A→ B work flow sequences. An item measuring this concept's complement, change induced in other subunits, confounded induced change and interaction; hence the analysis omits it. However, that item's correlation with interdependence was .58 as compared with an interitem correlation for interdependence of .64. Thus, we considered the measure to indicate general levels of policy and procedural interdependence. The four boundary-spanning activity items measured the frequency and time spent in interactions with members of other units.

Frequency of participation was measured for each of four decisions on a scale ranging from almost always to almost never. Two of the decisions were task-related and two concerned personnel. A fifth item concerned influence in changes in methods initiated by subordinates. Autonomy was a composite of two items reflecting independence of action generally and two items referring to independence in handling nonroutine decisions. Two of these specifically referred to independence from "supervisors or specialists" in making decisions. The measure confounded two aspects of autonomy: supervisory constraints and constraints placed upon members of focal units by other units. However, as Table 2 shows, interunit interdependence was not associated with autonomy ($r = -.08$), implying a bias toward supervisory constraints.

---

[2] Full texts of items and data for all analyses are available from the senior author.

Pooling

To compute scores for work groups, supervisors' scores were added to the means of subordinates' scores, and these sums were divided by two (Van de Ven et al., 1976). Three tests supported this procedure. One-way analysis of variance (Dixon & Massey, 1969) supported pooling subordinate responses into work groups for all of the variables ($p < .005$). No variable violated the homogeneity of variance assumption.[3] We dropped four units whose variances exceeded the pooled variance ($p < .05$) for two or more variables from the study.

Because there were scores for both supervisors and subordinates on each variable, a multitrait multimethod (MTMM) analysis (Campbell & Fiske, 1959) was conducted; it supported pooling supervisor and subordinate scores. First, the significance of the supervisor-subordinate correlations for each variable met the convergent validity criterion (Table 1). Second, if supervisor and subordinate scores were to be pooled, the pattern of relationships amongst the variables had to be similar in both heterotrait, monomethod blocks—the supervisor block and the subordinate block. Following Brief and Aldag (1978), we used Spearman rank order correlation to test the similarity of the two patterns. This yielded a correlation of .72 ($p < .05$) that supported pooling.

Methods Variance

Two tests assessed divergent validity. First, all 40 comparisons met the criterion that the convergent validity correlations ($r_{ii'}$) be greater than $r_{ij'}$ and $r_{i'j}$, where i and j denote supervisors' scores on variables i and j; similarly, i' and j' denote subordinates' scores on variables i and j. The second criterion, that $r_{ii'}$ be greater than r_{ij} and $r_{i'j'}$, was met in 31 of the 40 comparisons. Table 1 shows the nine exceptions. Interdependence and boundary-spanning activity show differing patterns with other variables (Table 2), thus mitigating concern about common methods variance. However, conclusions regarding autonomy bear a caveat on potential methods variance.

RESULTS AND DISCUSSION

Hypotheses 1 and 2 stated that interunit interdependence and task difficulty would be positively related to boundary-spanning activity. Results supported these hypotheses with $p < .001$ and $p < .01$, respectively (Table 2).

---

[3] Lowest $p = .30$ with the Bartlett - Box test.

TABLE 1
Results of Convergent and Divergent Validity Analyses[a]

Variables	Convergent Validity	Exceptions to Divergent Validity Criterion	
	Supervisor-Subordinate Correlations ($r_{ii'}$)	Supervisor Correlations ($r_{ii'} < r_{ij}$)	Subordinate Correlations ($r_{ii'} < r_{i'j'}$)
1. Task difficulty	.58***		
2. Interunit interdependence	.30**	r_{23} = .52***	$r_{2'3'}$ = .50***
3. Boundary-spanning activity	.50***	r_{32} = .52***	$r_{3'2'}$ = .50***
4. Participation in decision making	.48***		
5. Autonomy	.28**	r_{51} = .47*** r_{53} = .34** r_{54} = .32**	$r_{5'1'}$ = .32** $r_{5'4'}$ = .34**

TABLE 2
Descriptive Statistics and Zero-Order Correlations among All Variables

Variables	Unit Means[a]	Standard Deviations	Reliabilities[b]	Zero-Order Correlations among Variables[a]			
				2	3	4	5
1. Task difficulty	3.64	0.98	0.72	.03	.32**	.21*	.44***
2. Interunit interdependence	3.82	1.18	0.68		.51***	.09	−.08
3. Boundary-spanning activity	3.59	0.94	0.80			.38***	.30**
4. Participation in decision making	4.21	1.04	0.79				.28*
5. Autonomy	4.69	0.81	0.71				

[a] N = 71

[b] α; N = 269

*p < .05

**p < .01

***p < .001

These findings from an administrative setting complement Tushman's (1977, 1978) findings in a high technology setting. However, although Tushman found boundary-spanning activity to be concentrated in a few roles, the data in the present study did not permit an analysis of its distribution.

Hypothesis 3, that task difficulty would be positively related to participation, received only weak support ($r = .21, p < .05$), which is consistent with Mohr's (1971) and Hrebiniak's (1974) results. When task difficulty was correlated with each of the decisions, only two correlations were significant at $p < .05$; these were number of people and influence in subordinate-initiated changes in methods. Task difficulty was not related to two task decisions, introduction of new tasks and supervisor-initiated changes in methods. Other considerations, such as a need for acceptance of decisions (Vroom & Yetton, 1973), may overshadow task difficulty in predicting participation (Mohr, 1971). However, Duncan (1973) showed that linkage between uncertainty and participation was contingent upon the nonroutineness of decisions themselves, a characteristic this index did not measure. Therefore, this finding may reflect a flaw in our methods.

Hypothesis 4, linking task difficulty with autonomy, was supported at $p < .001$, consistent with studies cited earlier in this paper (e.g., Schoonhoven, 1981). In this study, autonomy was the dependent variable; however, Glisson (1978) argued for a linkage from autonomy to task difficulty, suggesting that workers given discretion may perceive variations in clients' needs and undertake nonroutine searches to solve problems. Nevertheless, in contrast to findings for participation, these results support Mohr's (1971) position that autonomy and participation be considered as separate variables, rather than be subsumed under a rubric such as "organicness" (Leifer & Huber, 1977:237). In this light, these variables make different contributions to information-processing capacity; therefore they may enjoy different relationships with variables that affect information-processing requirements.

Hypotheses 5 and 6, stating that boundary-spanning activity would be positively linked with participation and autonomy, received support at $p < .001$ and $p < .01$, respectively; these results were consistent with Leifer and Huber (1977). One contribution of this paper was isolating interunit boundary-spanning activity as affecting two dimensions of subunit structure, both of which are aspects of the influence of subordinates in decision making. Although our rationale included participation and autonomy as dependent variables, the relationships may be reversed or there may be mutual causation. Participation in decisions with extraunit implications may require boundary-spanning activity during implementation; where autonomy and interunit interdependence exist, unit members may be encouraged to seek out extraunit information through personal efforts. Longitudinal research to assess causal directions would facilitate understanding how changing designs to match requirements and capacity for information processing may stimulate shifts, or pressures for shifts, in the degree of influence subordinates exercise in decision making. Thus, such research would have implications for supervisory style. It would also help to reveal how shifts in subordinate influence may cause design changes to emerge.

This study has a number of limitations. First, as the Methods section indicated, there was potential common methods variance for autonomy. Second, only one organization, a people-processing state agency, was studied, and the units of analysis were work groups whose members had no supervisory responsibilities. Finally, this study used no measures of effectiveness; consequently we can draw no conclusions about whether the relationships found characterized effective or ineffective units.

CONCLUSIONS

Results support all six hypotheses and also support and extend research in information processing as an approach to understanding organizational design and structure. Two particular themes may contribute to further research in this area. The first is the need to specify design elements' particular contributions to information-processing capacity and relate these to specific demands for capacity that such factors as task difficulty induce. The second is the relationship between coordination mechanisms and elements of intraunit structure. Research in this area may assist practitioners by helping to predict changes in relationships between supervisors and subordinates that design changes may induce—or vice versa.

REFERENCES

Aiken, M., & Hage, J. 1968. Organizational interdependence and intraorganizational structure. ***American Journal of Sociology***, 33: 912–930.

Bacharach, S.B., & Aiken, M. 1977. Communication in administrative bureaucracies. ***Academy of Management Journal***, 20: 365–377.

Bell, G.D. 1967. Determinants of span of control. ***American Journal of Sociology***, 73: 90–101.

Brief, A.P., & Aldag, R.H. 1978. The job characteristic inventory: An examination. ***Academy of Management Journal***, 21: 659–670.

Burns, T., & Stalker, G.M. 1961. ***The management of innovation***. London: Tavistock Publications.

Campbell, D.T., & Fiske, D.W. 1959. Convergent and divergent validation by the multitrait-multimethod matrix. ***Psychological Bulletin***, 56: 81–105.

Cheng, J.L.C. 1983. Interdependence and coordination in organizations: A role-system analysis. ***Academy of Management Journal***, 26: 156–162.

Connolly, T. 1977. Information processing and decision making in organizations. In B.M. Staw & G.M. Salancik (Eds.), ***New directions in organizational behavior***: 205–234. Chicago: St. Clair Press.

Daft, R.L., & Macintosh, N.B. 1981. A tentative exploration into the amount and equivocality of information processing in organizational work units. ***Administrative Science Quarterly***, 26: 207–224.

Dixon, W., & Massey, F. 1969. ***Statistical analyses***. New York: McGraw-Hill Book Co.

Duncan, R. 1973. Multiple decision making structures in adapting to environmental uncertainty. ***Human Relations***, 26: 273–291.

Fry, L.W. 1982. Technology-structure research: Three critical issues. ***Academy of Management Journal***, 25: 532–553.

Galbraith, J.R. 1972. Organization design: An information processing view. In J.W. Lorsch & P.R. Lawrence (Eds.), ***Organization planning: Cases and concepts***: 49–74. Homewood, Ill.: Richard D. Irwin.

Galbraith, J.R. 1973. ***Designing complex organizations***. Reading, Mass.: Addison-Wesley.

Galbraith, J.R. 1977. ***Organization design***. Reading, Mass.: Addison-Wesley.

Gerwin, D. 1981. Relationships between structure and technology. In P.C. Nystrom & W.H. Starbuck (Eds.), ***Handbook of organizational design***, vol. 2: 3–38. Oxford: Oxford University Press.

Glisson, C.A. 1978. Dependence of technological routinization on structural variables in human service organizations. ***Administrative Science Quarterly***, 23: 383–395.

Grimes, A.J., Klein, S.M., & Schull, F.A. 1972. Matrix model: A selective empirical test. ***Academy of Management Journal***, 15: 9–31.

Hage, J., & Aiken, M. 1969. Routine technology, social structure, and organizational goals. ***Administrative Science Quarterly***, 14: 366–376.

Hrebiniak, L.B. 1974. Job technology, supervision, and work unit structure. ***Administrative Science Quarterly***, 19: 395–410.

Katz, D., & Kahn, R.L. 1978. ***The social psychology of organizations*** (2nd ed.). New York: John Wiley & Sons.

Lawrence, P.R., & Lorsch, J.W. 1967. ***Organization and environment***. Cambridge, Mass.: Harvard University Press.

Leifer, R., & Huber, G.P. 1977. Relations among perceived environmental uncertainty, organization structure, and boundary-spanning behavior. ***Administrative Science Quarterly***, 22: 235–247.

Litwak, E. 1961. Models of organization which permit conflict. ***Journal of Sociology***, 67: 177–184.

March, J., & Simon, H. 1958. ***Organizations***. New York: John Wiley & Sons.

McCann, J.E., & Ferry, D.L. 1979. An approach for assessing and managing inter-unit interdependence. ***Academy of Management Review***, 4: 113–120.

McCann, J.E., & Galbraith, J.R. 1981. Interdepartmental relations. In P.C. Nystrom & W.H. Starbuck (Eds.), ***Handbook of organizational design***, vol. 2: 60–84. Oxford: Oxford University Press.

Mohr, L.B. 1971. Organizational technology and organization structure. ***Administrative Science Quarterly***, 16: 444–459.

Pennings, J. 1975a. The relevance of the structural-contingency model for organizational effectiveness. ***Administrative Science Quarterly***, 20: 393–410.

Pennings, J. 1975b. Interdependence and complementarity—The case of a brokerage office. ***Human Relations***, 29: 825–840.

Perrow, C.A. 1967. Framework for the comparative analysis of organizations. ***American Sociological Review***, 32: 194–208.

Schoonhoven, C.B. 1981. Problems with contingency theory: Testing assumptions hidden within the language of contingency "theory." ***Administrative Science Quarterly***, 26: 349–377.

Thompson, J.D. 1967. ***Organizations in action***. Chicago: McGraw-Hill Book Co.

Thompson, J.D., & Bates, F.L. 1957. Technology, organization, and administration. ***Administrative Science Quarterly***, 2: 325–342.

Tushman, M.L. 1977. Special boundary roles in the innovation process. ***Administrative Science Quarterly***, 22: 587–605.

Tushman, M.L. 1978. Technical communication in R & D laboratories: The impact of project work characteristics. ***Academy of Management Journal***, 21: 624–645.

Tushman, M.L. 1979a. Work characteristics and subunit communication structure: A contingency analysis. ***Administrative Science Quarterly***, 24: 82–97.

Tushman, M.L. 1979b. Impact of perceived environmental variability on patterns of work related communication. ***Academy of Management Journal***, 22: 482–500.

Tushman, M.L., & Nadler, D.A. 1978. Information processing: An integrating concept in organization design. ***Academy of Management Review***, 3: 613–624.

Tushman, M.L., & Scanlan, T.J. 1981. Boundary spanning individuals: Their role in information transfer and their antecedents. ***Academy of Management Journal***, 24: 289–305.

Van de Ven, A.H., & Delbecq, A.L. 1974. A task contingent model of work unit structure. ***Administrative Science Quarterly***, 19: 183–197.

Van de Ven, A.H., Delbecq, A.L., & Koenig, K., Jr. 1976. Determinants of coordination modes within organizations. ***American Sociological Review***, 41: 322–338.

Vroom, V.H., & Yetton, P.W. 1973. ***Leadership and decision-making***. Pittsburgh: University of Pittsburgh Press.

Withey, M., Daft, R.L., & Cooper, W.H. 1983. Measures of Perrow's work unit technology: An empirical assessment and a new scale. ***Academy of Management Journal***, 26: 45–63.

Jack K. Ito is an associate professor in the Faculty of Administration at the University of Regina, Regina, Canada. His current research interests include processes affecting MBO systems.

Richard B. Peterson is Professor of Management and Organization at the University of Washington. His central research interests are in the areas of labor relations and international management. He is completing a study funded by the National Science Foundation on labor grievance procedures as well as continuing work on the changing Japanese employment system.

© *Academy of Management Journal*
1986, Vol. 29, No. 1, 149–159.

SOLUTION DIVERSITY AND THE QUALITY OF GROUP DECISIONS

JOHN P. WANOUS
Ohio State University
MARGARET A. YOUTZ
Michigan State University

Research investigating the effects of diversity on group decision making has been extensive. For example, Shaw's (1981) review discussed the effects

---

The authors would like to thank the following for their comments: Arnon E. Reichers, Richard A. Guzzo, Elaine D. Pulakos, John A. Wagner III.

of diverse needs, personality factors, abilities, genders, races, and other differences among individuals within groups. The presumed benefit of a diverse group is that its members' different points of view lead to a diverse set of recommended solutions. Researchers have suggested that a need to reconcile diverse solutions stimulates effective group discussion that leads to a high quality decision. Studies that measure solution diversity directly and relate it to the quality of group decisions are rarer than studies measuring diversity in group composition. Yet, investigators have considered solution diversity to be important. Examples are Maier's work on creative group decisions (Maier, 1963, 1967, 1970; Hoffman, Harburg, & Maier, 1962) and Janis's (1972, 1982) accounts of groupthink. Janis recommends a number of ways to prevent or overcome effects of groupthink; many of these involve increasing the diversity of solutions to problems that groups consider.

The present study continues to examine the link between solution diversity among a group's members and the quality of a group's final decisions. There are at least three reasons to continue research in this area. First, despite the popularity of both Maier's and Janis's advice that group leaders encourage solution diversity, relatively little supporting empirical research has emerged. Thus, research is needed that substantiates a direct link between solution diversity and the quality of group decision. If such a link can be established, systematic research can continue to investigate which interventions by group leaders (Janis, 1982; Maier & Hoffman, 1960, 1965) best stimulate solution diversity.

A second reason for conducting the present study was that much of the previous research lacked good criterion measures of the quality of group decisions. Janis's work is mostly narrative. Maier's measure of decision quality was an experimenter judgment, usually dichotomous, of group decisions.[1] Such a criterion is flawed to the extent that experimenters know the hypothesis being tested; moreover, researchers typically consider categorical measures less desirable than continuous measures. Furthermore, in Maier's research group, members acted out scripted roles, rather than simply being themselves; this, too, could contaminate the criterion.

A third reason to continue such research is that solution diversity is a form of group member diversity that can be more easily manipulated than other types. For example, it is generally not practical, and frequently impossible, for managers to sort organization members into small groups in terms of demographic, need, personality, or ability factors for particular problems, and then re-sort them for new problems. Thus, previous research on diversity in group composition[2] is of limited usefulness to most managers.

However, solution diversity is not a pure individual difference variable, as diversity among group members is. Rather, it is the result of interaction between people and situations. The amount of solution diversity within a given group of people varies with problems, but group composition factors stay constant across varied problem situations. Because solution diversity is not a fixed characteristic of groups or their members, it can vary naturally across problems. Group leaders can also consciously manipulate solution diversity along the lines Maier (1970) and Janis (1982) suggested; these include having a devil's advocate, inviting outsiders, or holding second chance meetings.

---

[1] See Maier (1970) for an entire research program.

[2] See Shaw (1981) for a review.

METHODS

Research Setting and Participants

Data were gathered from 98 ad hoc groups of students ($N = 594$) who were enrolled in an introductory course on management and organizational behavior. As part of this course students did several experiential exercises to illustrate behavioral science concepts. One of these exercises concerned differences between decisions made by individuals working alone and those made by groups. Data gathered in this classroom context do not neatly fall into either a laboratory or a field category, so it is important to specify how the present study compares to stereotypes of field and experimental lab studies.

The physical setting, a college classroom, was a more natural setting than the typical behavioral science lab in which to study group dynamics; however, it was a more temporary setting than is typical in work organizations. The motivation of the participants in the present study was to learn something about group dynamics and themselves. They did not know these data would be collected after the exercise and had no reason to see the exercise as a situation in which they were subjects for study. This probably reduced the artificiality of the experience and also reduced the likelihood of experimental demand characteristics being factors here. Gathering data from a learning exercise in a classroom also meant the behavior and the setting were consistent with each other.

Instructors asked their students to form small groups for the decision making exercise.[3] Students formed natural groups, typically joining those sitting in adjacent seats. Thus, groups were not random, as in a typical lab study. The process of group formation resembled that seen in friendship groups, or in informal groups in organizations like cliques and social groups.

As a result of the characteristics of the group formation process, groups varied in size from three to nine members; two-thirds of the groups had four, five, or six members. We removed the effects of group size and of odd or even numbers of members by hierarchical regression. Details appear in the Data Analysis section.

---

[3] Differences among the ten instructors could not affect these data because their role during the exercise was simply to read the instructions and hand out materials. However, after the groups had generated data, the instructors' role became more active. They read the expert opinions, and after groups calculated scores they conducted class discussions of the results. Therefore, instructor differences only appear after all the data have been generated.

Although this research setting was in some ways more natural than that of a typical experimental lab study, it was not equivalent to a field setting. First, the problem or task to be solved did not arise as part of a job, although it had some elements of a real world decision, as the Discussion section explains. Second, the groups were ad hoc rather than somewhat permanent work groups; they resembled temporary committees. Third, the consequences of individuals' behaviors were much less significant than would be the case in true organizational field settings.

Task

Two very similar group consensus ranking tasks were used: Subarctic Survival (Eady & Lafferty, 1974) and Mountain Survival (Kast & Rosenzweig, 1976). Each task has individuals assume they are stranded, either in the subarctic or on a mountain, and then has them rank order 15 items according to their importance for survival. First students ranked the items alone; then the groups were formed and instructed to reach group consensus on the ranking of the 15 survival items. The task instructions, which were read to the students, defined consensus as discussing a problem among themselves and trying to reach a group decision about the rank ordering of the 15 items. They were instructed not to vote to resolve conflict, nor to average their individual solutions to reach a group solution, because voting and averaging would defeat the purpose of the exercise. Finally, class instructors announced the correct rank orders, based on expert opinions provided by the developers of these tasks.[4] The scoring of individual and group solutions was then completed, as described in the next subsection.

Measures

The quality of group decisions was determined by comparing groups' rankings of the items to the expert ranking. Specifically, we determined the absolute value of the differences between a group's ranking of the 15 items and the expert ranking. The sum of these 15 absolute differences was the measure of decision quality. The lower this number, the higher the decision quality, because low numbers meant that a group's rank order closely matched that of the experts.

Five control variables were used to predict the quality of group decisions before we added the research variable of interest, solution diversity among group members. The five were: (1) group size; (2) odd or even number of group members; (3) type of group decision problem—subarctic or mountain survival; (4) best individual score; and (5) average of individual scores. We referred to these five variables as control variables in the hierarchical regression equation, because no experimental design controlled their effects. However, their entry into the regression equation prior to solution diversity statistically controlled these variables. We discuss our reasons for choosing these control variables and for their sequence in the regression in the Data Analysis section.

---

[4] The experts for the Subarctic Survival Problem were the Para-rescue Specialists for the 413th Transport and Reserve Squadron, the Canadian Forces, stationed at Summerside, Prince Edward Island. The squadron is responsible for air and sea rescue operations in Quebec, Newfoundland, Labrador, and the Arctic Regions. The expert opinion for the Mountain Survival Problem was provided by David Kneeland, Director of the Institute for Survival Education in Seattle, Washington.

The variable of major concern, solution diversity among group members, was measured with Kendall's coefficient of concordance, *W*. This statistic assesses the overall degree of agreement in a set of rankings given by several individuals (Siegel, 1956: 229–239). The coefficient of concordance is the n-case extension of a Spearman rank order correlation coefficient (r_S), or Kendall's rank order coefficient (tau). Values of *W* range from 0 to 1, not −1 to +1, as in a Pearson correlation, with low values indicating low agreement.

Data Analysis

The data were analyzed by hierarchical multiple regression in which the control variables were entered first into the regression equation in a specific temporal order. The measure of solution diversity among group members, Kendall's *W*, was entered last to determine how much of the remaining predictable variance in the quality of group decisions it accounted for. We used an *F*-test to calculate the statistical significance of changes in R^2 values at each step of the hierarchical ordering of variables.

The logic for ordering these variables in the regression equation was as follows: We entered total group size first because size could affect the level of solution diversity; however, the reverse is illogical. Second, whether the group had an odd or even number of members was entered before solution diversity because it too could affect levels of diversity. We considered this a possibility because voting to reduce conflict is more likely with odd numbers of group members. Third, we entered the type of ranking task before solution diversity, using a dummy code, because there could be differences between the two tasks used here.

After statistically controlling for group size, odd-even composition, and problem type, we entered the two measures of individual solution quality—best individual score and the average of individual scores—as the fourth and fifth control variables. These were entered before solution diversity because they represented the temporal starting points of group discussions. Furthermore, both measures of individual solution quality also acted as controls for two critical individual differences, general IQ and specific expertise, that normally would be controlled for by randomization of subjects to groups.

RESULTS

Table 1 shows the means, standard deviations, and intercorrelations among the variables. The pattern of intercorrelations confirms the necessity of entering the five control variables into the regression equation prior to entering solution diversity. First, as group size increases, the level of diversity (*W*) also increases ($r = -.36$, $p < .01$), as had been anticipated. Second,

TABLE 1
Means, Standard Deviations, and Intercorrelations of Variables

Variables	Means	s.d.	1	2	3	4	5	6
1. Group size	6.01	1.51						
2. Odd or even number of group members[a]	0.50	0.50	.06					
3. Problem type[b]	0.55	0.50	−.58**	−.10				
4. Best individual score[c]	41.77	9.14	−.08	.02	−.08			
5. Average of individual scores[c]	57.14	6.39	.22*	.11	−.42**	.63**		
6. Solution diversity[d]	0.56	0.10	−.36*	−.11	.30**	.08	−.27	
7. Decision quality[c]	43.56	11.43	.12	.10	−.27**	.52**	.66**	.03

[a] Coded 0 = odd number of group members, 1 = even number of group members.
[b] Coded 0 = Mountain Survival Task, 1 = Desert Survival Task.
[c] Low values mean high scores or high decision quality.
[d] Low values mean high diversity.
*$p < .05$
**$p < .01$

there was a slight tendency ($r = -.11$, n. s.) for increasing diversity to be found in groups with even numbers of members, as expected. Third, we found removal of the variance owing to the type of task to be desirable, since the type of task was correlated with both solution diversity ($r = .30$, $p <.01$) and quality of group decision ($r = -.27$, $p < .01$). The Subarctic Survival task produced less diversity and was easier than the Mountain Survival task. Finally, removing the effects of the best individual score and the average of the individual scores was desirable, because both were strongly correlated with group decision quality ($r = .52$, $p < .01$ for best score, and $r = .66$, $p < .01$ for average score), although neither correlated significantly with solution diversity.

Table 2 shows the results of a hierarchical regression used to determine the incremental variance in the quality of group decisions accounted for by (1) the control variables and (2) solution diversity among group members. Examination of Table 2 indicates that the first variable, group size, did not account for a significant change in R^2, nor did the odd-even composition of a group. However, the type of problem did account for a significant amount of variance (6%, $p <.01$) in groups' decision quality scores. The best individual score accounted for 25 percent ($p < .01$) of the variance in group decision quality, and the average individual score accounted for an additional 13 percent ($p < .01$).

The degree of solution diversity among group members also accounted for a significant amount of variance (3%, $p < .05$) in group decision quality. Besides accounting for a significant portion of variance in the quality of group decisions, the effects of solution diversity are also in the direction expected. That is, the positive correlation in Table 1 and the positive beta

TABLE 2
Summary of Results of Hierarchical Regression Predicting Decision Quality

Variables	Cumulative R^2s	R^2 Changes	Betas
Group size	.01	.01	.046
Odd or even number of group members	.02	.01	.047
Problem type	.08	.06**	−.033
Best individual score	.33	.25**	.116
Average of individual scores	.46	.13**	.618**
Solution diversity	.49	.03*	.215*

*p < .05
**p < .01

weight ($p < .05$) in Table 2 show that increasing diversity is associated with increasing quality.

DISCUSSION

The results of this study provide support for the general advice that diversity in groups enhances decision quality. Among those studies that have specifically examined solution diversity, the present research demonstrates the effects of such diversity more precisely than earlier efforts. Given the procedures for data gathering, measurement of key variables, and conservative analysis method, these data show solution diversity to have a specific and identifiable positive effect on decision quality, even after the influence of several other important factors has been removed.

The fact that measures of individual solution quality were much stronger predictors of the quality of group decisions than solution diversity does not really weaken the conclusion that solution diversity enhances decision quality. The relationship between individual and group decision quality should be quite high for two obvious reasons. First, they are measures of the same type—quality—and thus share much common method variance. Second, doing the same task twice within about one hour's time[5] almost amounts to an immediate test-retest measure, which certainly would inflate a relationship.

Besides the two methodological reasons discussed above, there are other grounds for discounting the practical significance of the two individual solution quality scores, even though they account for most of the variance and are highly significant. Specifically, what can be said to group leaders other than "give the problem to your best subordinate, or if you don't know who that would be, make sure your group is of high average competence"? The first part of the advice assumes that leaders can identify such individuals; but frequently leaders cannot know a priori who will be the most competent. In fact, one reason why leaders use groups is because they cannot specify such individuals; this is particularly true when problems are unstructured (Vroom & Yetton, 1973). The second part of the advice is also a bit glib, since it presumes that leaders have enough control over the composition of work groups to ensure a high average level of competence. In contrast, managing solution diversity among group members appears to be something most managers can control.

---

[5] Groups' consensus solutions followed individuals' rank order solutions.

Other forms of diversity that have been studied, such as needs, personalities, abilities, or demographic factors, would be much more difficult to affect within a group in any reasonable period of time. The major implication for action from studies of diversity among group members has been to recommend appropriate composition and recomposition of groups. Practically speaking, this is quite difficult even in the most fluid organizations. Hence, Maier's (1963) recommendation, that leaders encourage groups to develop second solutions, or Janis's (1982) advice, that leaders appoint devil's advocates, invite outsiders to meetings, hold second chance meetings, and so forth, have greater relevance for actual managerial practice than recommendations of group recomposition.

With all the emphasis in the preceding section on the management of solution diversity, it must be remembered that the present research did not involve manipulations or interventions. Rather, the degree of naturally occurring solution diversity that members brought to the group discussion was related to decision quality. Specifying that diversity enhances decision quality is a necessary first step in research of this type. A subsequent study could investigate links between different types of interventions by group leaders, such as those mentioned in the preceding paragraph, and degrees of solution diversity generated as a result. Whether there might be a difference between naturally occurring solution diversity brought to meetings and that resulting from a leader's actions during meetings is also a potentially interesting question.

The present study has several limitations. First, the participants in this research were not well acquainted with one another, and they formed temporary groups solely for the purpose of this study. However, previous research found that the beneficial effects of solution diversity are strongest in established groups (Hall & Williams, 1966), because controversy is usually less threatening than it is in new groups. Because solution diversity did enhance the quality of group decisions in the present study using ad hoc groups, the effect of solution diversity found here could be a lower bound. Future research might profitably examine the effects of diversity on established versus ad hoc groups in order to estimate the upper bound of the effect as well.

A second limitation concerns the range of solution diversity, because extreme values of Kendall's coefficient of concordance did not occur here (mean $W = .56$, s.d. $= .10$). Whether extremely high or low solution diversity would continue to enhance decision quality is also a question for further research.

A third limitation is that solution diversity seems to enhance decision quality, but may lower decision acceptance and group member satisfaction

(Castore, 1973; Castore & Murnighan, 1978; Shaw, 1981). Since both quality and acceptance are components of overall decision-making effectiveness (Maier, 1963; Vroom & Yetton, 1973), trade-offs may need to be considered between them, particularly when a problem or decision has both high quality and high acceptance requirements.

The final limitation is that we were not able to obtain data about the process of group decisions in this study. Thus, there are no direct data, such as observations or self reports, showing precisely how solution diversity enhanced decision quality. Unsystematic personal observations of group process by the researchers and the instructors who helped gather these data suggest that diversity may increase decision quality by increasing the amount of attention and discussion paid to each group member's individual solution. The resulting scrutiny could easily lead to the surfacing of unrecognized and incorrect assumptions among group members. For example, consider this situation of low diversity: one group member ranks an item in the survival task eighth and a second member ranks it tenth. In such a case it will be tempting for the members to settle their minor conflict easily. In contrast, when one member ranks a particular item first that another ranks as fifteenth, there is an extremely high diversity. In this case, the members are much more likely to discuss the reasons for their rankings than in the case of low diversity. As a result, the assumptions each has made are more likely to surface, deepening their understanding of how to cope with the task.

The final point for discussion concerns external validity. Maier (1963) distinguished between solution generation and solution evaluation. It is clear that this study concerns the second of these two phases of decision making. In this sense, it can complement research on the process of solution generation, like brainstorming research. The fact that the present research addressed the solution evaluation stage exclusively might be assumed to be a major limitation to its external validity. However, this assumption is false. Highly structured problems with clear alternatives often arise, and the only problem is to evaluate them.

In general, any situation in which decision alternatives come to a group, rather than being generated by a group, resembles the task used in the present research. In such situations solution evaluation is the critical process, since it is under group control. An extremely common, analogous situation occurs when groups must evaluate job applicants and then rank order them so job offers can be extended. Similar examples include groups' deciding on the rank order of employees for promotions as positions open up, or ranking who should be sent for advanced training. In all cases the groups making the decisions have clearly defined alternatives and must decide on rank orders. Human judgment assesses the quality of a group's decision, as in the present case. Since these examples of personnel decisions are common to most organizations, our results should have fairly wide applicability.

These results may have even wider applicability if solution diversity enhances decision quality regardless of the particular type of problem. We gave a conservative view of external validity earlier in this section, stating that the results generalize only to structured problems concerned mainly with solution evaluation, because that situation mirrors the task used to generate the data. However, the works of both Maier (1970) and Janis (1982) concern tasks and problems with unstructured characteristics; their results and recommendations also support the beneficial effects of solution diversity on the quality of group decisions.

As a final note, there are four circumstances in which solution diversity may be counterproductive. Although none of these characterized the present study, they may be boundaries to the beneficial effects of solution diversity. First, if diversity leads to interpersonal conflict, overall group effectiveness may decline (Evan, 1965). Second, if diversity prevents agreement on basic goals for a group, paralysis may result. Third, in crises with severe time pressure, diversity may slow a group too much. Fourth, if a decision to be made has minimal quality requirements, but high acceptance requirements (Maier, 1963), solution diversity may prevent consensus.

REFERENCES

Castore, C.H. 1973. Diversity of group member preferences and commitment to group decisions. ***Annals of the New York Academy of Science,*** 219: 125–136.

Castore, C.H., & Murnighan, J.K. 1978. Determinants of support for group decisions. ***Organizational Behavior and Human Performance,*** 22: 75–92.

Eady, P.M., & Lafferty, J.C. 1974. ***The subarctic survival problem.*** Plymouth, Mich.: Experiential Learning Methods.

Evan, W.M. 1965. Conflict and performance in R&D organizations. ***Industrial Management Review,*** 7: 37–46.

Hall, J., & Williams, M.S. 1966. A comparison of decision-making performances in established and ad hoc groups. ***Journal of Personality and Social Psychology,*** 3: 214–222.

Hoffman, L.R., Harburg, E., & Maier, N.R.F. 1962. Differences and disagreement as factors in creative group problem-solving. ***Journal of Abnormal and Social Psychology,*** 64: 206–214.

Janis, I.L. 1972. ***Victims of groupthink.*** Boston: Houghton-Mifflin.

Janis, I.L. 1982. ***Groupthink*** (2nd ed.). Boston: Houghton-Mifflin.

Kast, F.E., & Rosenzweig, J.E. 1976. ***Experiential exercises and cases in management:*** 187–199. New York: McGraw-Hill Book Co.

Maier, N.R.F. 1963. ***Problem-solving discussions and conferences.*** New York: McGraw-Hill Book Co.

Maier, N.R.F. 1967. Assets and liabilities in group problem solving: The need for an integrative function. ***Psychological Review,*** 74: 239–249.

Maier, N.R.F. (Ed.) 1970. ***Problem solving and creativity.*** Belmont, Calif.: Brooks/Cole.

Maier, N.R.F., & Hoffman, L.R. 1960. Quality of first and second solutions in group problem solving. ***Journal of Applied Psychology,*** 44: 278–283.

Maier, N.R.F., & Hoffman, L.R. 1965. Acceptance and quality of solutions as related to leaders' attitudes toward disagreement in group problem solving. ***Journal of Applied Behavioral Science,*** 1: 373–386.

Shaw, M.E. 1981. ***Group dynamics: The psychology of small group behavior*** (3rd ed.). New York: McGraw-Hill Book Co.

Siegel, S. 1956. ***Nonparametric statistics for the behavioral sciences.*** New York: McGraw-Hill Book Co.

Vroom, V.H., & Yetton, P.W. 1973. ***Leadership and decision making.*** Pittsburgh: University of Pittsburgh Press.

John P. Wanous, who earned his Ph.D. degree at Yale University, is Professor of Management and Human Resources and Joint Professor of Psychology at Ohio State University. He was previously on the faculties of Michigan State and New York University. His research has concerned work motivation and job satisfaction, with a particular emphasis on the entry of newcomers to organizations. His current research interests focus on realistic job previews and on the relationship between newcomers and the work groups they join.

Margaret A. Youtz is a doctoral student in the Department of Management at Michigan State University. Her general research interests concern personnel and microorganizational behavior, specifically supervisor performance ratings and how people choose jobs.

© *Academy of Management Journal*
1986, Vol. 29, No. 1, 159–169.

USING ELECTRONIC MAIL FOR DATA COLLECTION IN ORGANIZATIONAL RESEARCH

LEE S. SPROULL
Carnegie-Mellon University

Data for organizational research are often collected directly from organizational members, either through paper-and-pencil questionnaires or oral interviews. For example, of the 60 articles published in the *Academy of Management Journal* in 1983, 34 were based on data collected by questionnaire or interview. (See the Appendix.) Although questionnaires and interviews differ in many subtle and important ways, two of the most powerful differences concern costs and response rates. Interviews are generally more costly in researcher and respondent time than are questionnaires, but they yield a higher response rate on a comparatively smaller sample size (Kidder, 1981: 148–155). Every organizational researcher faces a cost/response tradeoff when choosing data collection methods. This paper describes a new tool for data collection—electronic mail (EMS)—and suggests that under certain conditions it may produce higher response rates at a lower cost than either questionnaires or interviews.

The data reported in this paper were collected while the author was an NIMH Research Fellow at Stanford University. A grant from the System Development Foundation supported preparation of the manuscript. Jonathan Cender, Martha Feldman, Sara Kiesler, James March, and Bob Sproull contributed substantially to the ideas and analysis presented here—the interesting parts anyway.

Electronic mail uses computer text-editing and communications tools to provide a high speed message service. Anyone with a computer account can use a terminal to compose a message and send it to anyone who has a mailbox on that computer or on any other computer that communicates with the sending computer. Four important characteristics of EMS are its speed, asynchrony, lack of intermediaries, and ephemerality of messages. Messages can be transmitted in seconds. They are sent at the convenience of senders and read at the convenience of recipients. The scheduling constraints of synchronous communication vanish. Senders and receivers usually process their own mail; they do not submit messages to or receive messages from secretaries or operators who process messages for many people. Messages appear on and disappear from video screens with no hard copies left behind. Every system makes it possible to store messages on computer files and to create hard copies of them, but most messages are never translated into hard-copy form. These four characteristics lead to a communication system that is relatively fast and convenient to use. Organizations currently using EMS include AT&T, Bank of America, Digital Equipment Corporation, Hewlett-Packard, IBM, Manufacturer's-Hanover Trust, Westinghouse, and 3-M.

As EMS becomes more prevalent in organizations, it may become an appropriate medium for collecting data from organization members. For example, a researcher could easily send out questionnaires via electronic mail. Recipients would simply edit in their responses to questions and then return edited questionnaire messages to the original sender. It would seem that researchers could collect data rapidly and conveniently this way. Postal mail delays would not occur. Respondents could fill out questionnaires at their convenience.

Three general issues should be considered in exploring the feasibility of using EMS as a research tool: (1) respondents' access to EMS; (2) their willingness to respond to surveys via EMS; and (3) the comparability of data collected via EMS with those collected by more conventional means. Although use of electronic mail is increasing rapidly, it currently reaches and is used by only small numbers of organizational members. When EMS users are an appropriate research population for particular research projects, accessibility is not a problem. As EMS use increases, the issue of access will decrease in significance. A second issue is willingness of potential respondents to participate in research conducted via EMS. If potential respondents see such use of EMS as illegitimate, aversive, or unrewarding, participation rates may be low in comparison with those for other media. Respondents may worry about possible violations of privacy or lack of anonymity. They may find the medium too impersonal for some kinds of questions. Evidence suggests that organizational members enjoy using EMS in their jobs and evaluate it quite positively (Crawford, 1982; Nyce & Groppa, 1983). This does not mean they will necessarily respond enthusiastically to EMS surveys, but it does suggest that they will not be discouraged by aversion to the medium itself. The third issue is substantive differences in answers that might be produced as a result of communicating via EMS versus some other medium. There is a tradition in social science research of trying to understand how the means of collecting data affects that which is collected (e. g., Stewart & Cash, 1978; Sudman & Bradburn, 1982). If EMS for some reason produces data that differ from those generated by other media, its use could not be justified without comparative research on the nature of those differences.

This paper investigates the second and third issues. It describes an organizational research project carried out entirely by electronic mail. It contrasts the procedures and results of this project with (1) those of a comparison project conducted in the same organization using conventional data collection techniques, and (2) with response rates reported in the literature.

RESEARCH SAMPLE AND PROCEDURE

The R&D and product development divisions of a *Fortune* 500 office products and systems manufacturer were the setting for this research project. Within each division one unit was selected that represented a well-established electronic mail community. This organization has used EMS for 12 years; over 80 percent of all employees in the units selected have excellent access to EMS terminals. They use the system extensively; the mean message traffic per day per person is 24 messages. All participants in this study were professionals. Ten men were selected at random from the professional ranks of the research organization; ten men and ten women were selected from the professional ranks of the product development organization. Thus, the size of the total target sample was 30 professionals from well-established electronic mail communities. The comparison project had a comparable sample of 30 randomly selected professionals. Respondents in the two samples did not differ on age, education, job category, number of years in the organization, number of years using EMS, or access to EMS. Their mean age was 34; their mean number of years of schooling was 16; their mean number of years in the organization was 4; their mean number of years using EMS was 5;[1] and 78 percent had EMS terminals in their offices.

The general data collection protocol asked participants to save all of their electronic messages for three days and to then answer a series of questions about some of them. Participants were asked 10 open-ended questions about the nature of each message, for example,"Why did you receive (or send) this message? What is it about?" They were asked 13 closed-form questions about each message partner, for example, "What is this person's hierarchical position with respect to you?" with 1 = subordinate, 2 = peer, 3 = superior, 4 = not in chain of command. Each participant also filled out an 80-90 item questionnaire about general EMS attitudes and behaviors. The questionnaire included items on personal history, EMS history, evaluation of current EMS programs, self-reports of EMS behavior, and questions about how respondents would choose to do a series of hypothetical tasks.

---

[1] Some respondents had used EMS in other organizations in which they had been employed.

In the EMS version of this research, all communication was via electronic mail. The researcher, who had a mail account on one of the corporation's computers, sent an invitation to participate in the project and an endorsement from the division vice-president as an electronic message. The invitation message asked participants to send back messages indicating interest. When a participant responded, the researcher sent a message explaining the data collection procedure. This message told participants to save all of their messages for three days in a computer file and then send the file to the researcher, who would send back a sample of messages with questions to be answered about each. After a participant returned three days' worth of messages, the researchers sent back all messages from the third day, or a random sample of 15 if there were more than 15 messages on that day. Each went as a separate message from the researcher with this statement appended:

> Thank you for sending me your messages. I have chosen a random sample of 15 of them [or, a sample of them] about which to ask you some questions. I am sending each of these messages back to you with questions appended on the end. Please use the "forward" command to copy this message into the edit window, then edit your answers in to each question, then "deliver" the result back to me. I am also sending, as the 16th message, a brief questionnaire; you should also forward it back to me with your answers edited in.

Ten open-ended questions about the nature of the message and 13 closed-form questions about characteristics of message partners' followed each message. The final message sent a questionnaire about general EMS attitudes and behaviors that included nine items about the experience of providing data electronically.

The conventional study was designed to produce a very high response rate. Its participants received a general letter explaining the nature of the research accompanied by an endorsement letter from the division vice-president. The interviewer then telephoned participants to answer any questions, explain the research procedure, and schedule interviews. They were told to save all of their messages for a three-day period and to have hard copies available at the time of their interviews. In face-to-face interviews they (1) orally answered 10 open-ended questions about all messages from the previous work day, or a random-sample of 15 if there were more than 15 messages on that day; (2) filled out a 13-item closed-form characteristics sheet for each message partner; and (3) filled out a written questionnaire.

In summary, both groups were asked to save copies of their messages and make them available to the researcher either in an on-line file or a hard copy. Both groups answered the same open-ended questions about a sample of their messages, either electronically or orally. Both groups answered the same questions about their message partners, either electronically or in writing. Both groups filled out a general questionnaire either electronically or in writing. Participants in the electronic version also answered nine questions at the end of the questionnaire about the experience of providing data electronically. Participants in the conventional version were asked no comparable questions.

RESULTS

Participation

The participation rate in the electronic version of the study was 73 percent. The participation rate was 87 percent in the conventional study ($t = .86$, $p = $ n.s.). In a follow-up phone call, four out of six nonparticipants in the electronic version indicated they would have been willing to participate but they had forgotten to respond to the original message because they were too busy. Table 1 shows nonparticipation rates for the two studies.

The 73 percent participation rate is high in comparison with typical response rates in hard-copy mail questionnaires, which typically range from 20 to 50 percent (Heberlein & Baumgartner, 1978). Organizational researchers often do better than public opinion researchers, but response rates still are low enough that nonresponse bias can be a problem. For example, the mean response rate across the 14 mailed questionnaire studies published in the *Academy of Management Journal* in 1983 was 53 percent. (See the Appendix.)

A mail study conducted at the same time as the EMS study reported in this paper offers a direct, although unplanned, comparison of EMS with hard-copy mailed questionnaires. Steinfield (1983) conducted a study of 400 people who worked in the same divisions of the corporation as the participants in the study reported here. His study had the same corporate sponsor-

TABLE 1
Nonparticipation Rate for Two Different Means of Data Collection in a Field Research Project

	Means of Data Collection						
	EMS			Face-to-Face			
Reason for Nonparticipation	Research Men	Product Development Men	Product Development Women	Research Men	Product Development Men	Product Development Women	Totals
Did not respond	2	2	1	0	0	0	5
Refused	0	1	0	1	0	0	2
No longer employed	1	0	0	0	2	1	4
No EMS access	1	0	0	0	0	0	1
Total nonparticipants	4	3	1	1	2	1	12
Total participants	6	7	9	9	8	9	48
N=	10	10	10	10	10	10	60

ship as this one and included respondents from the same professional levels. The topic of that study was also similar to the topic of this one. He reported a response rate of 55 percent using hard-copy mailed questionnaires.

Duration of Data Collection

The average elapsed time for data collection was half as long in the electronic version as it was in the conventional version. In the face-to-face research, it was necessary to wait four days after mailing invitation letters before telephoning to schedule interviews to make sure that respondents had received their initial letters in the mail. Only about 60 percent of the telephone calls actually reached the participants on the first day, requiring callbacks on subsequent days. Interviews were scheduled a minimum of four days after phone calls so that participants could save three days worth of messages. The average elapsed time between initial letter and completed protocol in the face-to-face research was 12 days; the average elapsed time in the electronic mail version was 5.6 days ($t=9.3$, $p < .01$).

The time each respondent actually spent answering questions seemed to be about 15 minutes longer in the EMS version than in the face-to-face version. (EMS respondents received nine more questions than the face-to-face respondents.) The amount of time required for a face-to-face interview averaged just under one hour. The average time reported by participants in the EMS version for responding to all the messages was 74 minutes ($t=1.38$, p=n.s.).

Completeness of Data

The questionnaire data were slightly more complete in the conventional study than in the EMS study, although in both cases they were well within acceptable bounds. In the face-to-face version, the interviewer could check the questionnaires on the spot to make sure they had been filled out properly and completely. This was impossible in the EMS version, hence the missing data rate was higher. In the face-to-face version the overall mean missing data respondent was .2 percent; in the EMS version it was 1.4 percent ($t=1.16$, p=n.s.).

Substantive Comparisons

There were no differences in the nature of answers provided by EMS participants compared to those provided in the conventional study, as assessed by t-statistics. Both sets of participants reported the same amounts of message traffic. They reported the same EMS writing styles in terms of how often they used salutations and closings and fixed typos. They also reported the same EMS histories, the same frequencies of work and nonwork messages, the same incidence of inappropriate messages, and the same choices of medium for a series of hypothetical tasks.

Most of the questions in this study were factual or objective. We would not expect the medium through which they were asked and answered to affect the substance of people's responses. When questions tap subjective states such as emotions or definition of self, however, the medium used may be more likely to affect the nature of the answers. The stronger the social presence of the researcher, the more likely people are to report positive or socially desirable responses (Locander, Sudman, & Bradburn, 1976; Woods & McNamara, 1980). Other research has demonstrated that electronic responses, in comparison with oral or hard-copy responses, tend to be more extreme, less socially desirable, and characterized by more self-disclosure (Kiesler, Siegel, & McGuire, 1984; Kiesler & Sproull, 1986; Kiesler, Zubrow, Moses, & Geller, 1984). This is because electronic mail provides almost no information about the social presence of message partners or the social context of communications. Without the regulating effects of such social information, responses tend to be more extreme. The most subjective questions in this study were those about how people would choose to do 12 hypothetical tasks. As reported earlier in this section, there were no differences in mean responses to these questions across the two studies. There was, however, a difference in extremity of response. Each hypothetical question could be answered on a scale from 0 to 10; every 0 or 10 response was coded as an extreme response. The EMS study elicited more extreme responses than did the conventional study. Out of 48 hypothetical task questions, the EMS study had a larger number of extreme responses in 35 of them ($t = 3.24$, $p < .01$). This was the only way in which the two studies differed substantively. Extreme responses are important substantively, because they can indicate very strong opinions. The differences in extremity of response are noteworthy in this study because they suggest that there may be an interaction between method of data collection and question type for some questions but not others: method may affect subjective indicators but not objective or factual ones.

Attitudes toward Data Collection

Participants in the electronic mail version varied in their attitudes toward providing data via electronic mail, although mean attitudes were fairly neutral. Participants in the face-to-face version were not asked a comparable set of questions. Table 2 summarizes responses to a series of questions about the process of providing data via electronic mail.

It is possible to compare attitudes toward participating in further research using the same media as these studies—either electronic mail or face-to-face interviews. Participants in the face-to-face version were asked: "With your permission may I contact you again to talk more about electronic mail?" All of the face-to-face participants said yes. Participants in the EMS version were asked: "Would you be willing to be contacted again via electronic mail at some later date to answer more questions about your use of electronic mail?" Ninety percent of the EMS participants said yes. The attitudinal data, although certainly not conclusive, suggest no serious disadvantage for electronic mail in comparison with face-to-face data collection. Table 2 summarizes the attitudinal data.

TABLE 2
Distribution of Attitudes toward Providing Data via EMS

	Frequencies of Responses					
Attitudes[a]	1	2	3	4	5	Means
Convenience	4	9	4	5	0	2.4
Ease of providing data	3	5	8	6	0	2.8
How time-consuming?	0	4	8	4	6	3.5
How interesting?	2	3	11	3	3	3.1
Prefer EMS?	6	5	2	3	6	2.9

[a] All attitudes were measured on 5-point Likert scales. Questions were of the form "How convenient did you find this data collection process?" 1 = very convenient; 5 = very inconvenient.

DISCUSSION

In this study, an EMS questionnaire produced (1) a response rate of 73 percent; (2) an elapsed time for data collection of less than a week; (3) a missing data rate of 1 percent; (4) no substantive differences in mean responses to questions in comparison to a written questionnaire; (5) fairly neutral attitudes toward providing data via EMS; and (6) willingness to provide data again via EMS.

The population of this study included only professional men and women. Other kinds of employees—secretaries or managers, for example—should be equally able to provide data via EMS, so long as they use EMS naturally in their jobs. But of course researchers should investigate this issue empirically. The topic of this study was behavior relative to EMS. There is no reason to believe that other topics unrelated to EMS, like job characteristics or quality of work life, would lead to results different from those in this study. Clinical researchers, for example, have used on-line questionnaires to elicit reliable data about medical behaviors (Erdman, Klein, & Greist, 1983). But obviously different topics require further research. The questions in this study were relatively objective ones. Subjective indicators collected via EMS may produce responses that differ from those collected via conventional means. This issue is the topic of current research (Kiesler & Sproull, 1986).

Any researchers with access to their respondents' electronic mail system could easily follow the EMS data collection procedure described in this paper. The results of this study do suggest some possible modifications, however. The behaviors of nonparticipants and cautious participants in the EMS version indicate that participation was at least slightly more problematic for the electronic mail research than it was for the face-to-face research. There are two likely reasons for this finding. One is that electronic mail suppresses status cues to legitimacy. In the face-to-face version, the initial invitation appeared on letterhead stationery from the Stanford University Business School; the endorsement letter was written on a corporate memorandum form. Each letter was personally signed. These communications may carry more probative information about the status and legitimacy of their senders than do identically worded messages sent electronically. Perhaps formal hard-copy letters could precede EMS questionnaires. The second reason has to do with managing attention. In the electronic mail version, the research request was undifferentiated from the 20 or so other messages each person got during the day. The participants in this research get many fewer first-class letters than that; hence that form of communication is likely to command more attention, even apart from status cues. But additionally, in the face-to-face version, the researcher took the initiative and tracked down participants by telephoning as often as was necessary to actually talk with them. In the electronic mail version, each participant had to take the initiative to respond. Sending out electronic reminders could help here.

It is also possible to suggest some reasons why the electronic mail version of answering questions takes noticeably longer than the face-to-face version. One is simple differences in transcription rates. Most people can probably speak[2] faster than they can type. A second is pacing differences. In the face-to-face version the researcher did not want to take any more of the respondents' time than was necessary and therefore kept respondents answering rapidly. In the electronic mail version, no external goad kept up the pace. A third is differences in developing a shared shorthand. In the face-to-face version it was common for respondents to describe a message as "just like the previous one," thereby eliminating the need to ask every question about the message. Additionally, after a few messages, many respondents learned the nature and order of questions and were able to answer them for each new message without prompting. The noninteractive nature of the electronic mail version prohibited the development of such shared understandings.

EMS may prove to be a quite useful tool for organizational researchers. Two of its important strengths, demonstrated in this paper, are convenience and low cost. Transmission time is low in comparison with mailed questionnaires. Respondents reply at their own convenience, a feature that may outweigh the fact that responding to an electronic questionnaire may take slightly longer in elapsed time than would a comparable face-to-face or hard-copy questionnaire. Researchers need not be physically present to deliver or administer questionnaires. One potential weakness is that there is no social presence creating a pressure to respond, so response rates may be lower than in face-to-face interviews—but perhaps higher than with hard-copy mailed questionnaires. Another weakness is that it is difficult to convey pictures or other graphical information with current EMS technology. In the study reported in this paper, EMS was clearly the preferred data-collection method because it produced adequate data, response rates, and willingness for further participation, with little expenditure of researcher time or effort and a high degree of convenience for respondents.

---

[2] In this study, the researcher took summary notes by hand.

REFERENCES

Crawford, A. B. 1982. Corporate electronic mail—A communication-intensive application of information technology. ***MIS Quarterly***, 6: 1–14.

Erdman, H., Klein, M., & Greist, J. 1983. The reliability of a computer interview for drug use/abuse information. ***Behavior Research Methods and Instrumentation***, 15: 66–68.

Heberlein, T., & Baumgartner, R. 1978. Factors affecting response rates to mailed questionnaires. ***American Sociological Review***, 43: 447–462.

Kidder, L. H. 1981. ***Research methods in social relations*** (4th ed.). New York: Holt, Rinehart & Winston.

Kiesler, S., Siegel, J., & McGuire, T. 1984. Social psychological aspects of computer-mediated communications. ***American Psychologist***, 39: 1123–1134.

Kiesler, S., & Sproull, L. 1986. Response effects in the electronic survey. ***Public Opinion Quarterly***, forthcoming.

Kiesler, S., Zubrow, D., Moses, A. M., & Geller, V. 1985. Affect in computer-mediated communication. ***Human Computer Interaction***, 1: 77–104.

Locander, W., Sudman, S., & Bradburn, N. 1976. An investigation of interview method, threat and response distortion. ***Journal of the American Statistical Association***, 71: 269–275.

Nyce, H. E., & Groppa, R. 1983. Electronic mail at MHT. ***Management Technology***, 1(1): 65–72.

Steinfield, C. 1983. ***Communicating via electronic mail: Patterns and predictors of use in organizations***. Unpublished doctoral dissertation, University of Southern California, Los Angeles.

Stewart, C. J., & Cash, W. B., Jr. 1978. ***Interviewing: Principles and practices***. Dubuque, Iowa: William C. Brown.

Sudman, S., & Bradburn, N. 1982. ***Asking questions***. San Francisco: Jossey-Bass.

Woods, K., & McNamara, J. R. 1980. Confidentiality: Its effect on interviewee behavior. ***Professional Psychology***, 11: 714–720.

APPENDIX

For every article and research note published in the *Academy of Management Journal* in 1983 the means of data collection was noted. Out of 60 articles, 34 used some form of questionnaire or interview. For each of these articles, three additional characteristics were noted: the target sample size, the final sample size, and the response rate (final divided by target). In some cases, one or more of these pieces of information did not appear but could be inferred from Table Ns or calculated from other information. Table A.1 summarizes the characteristics of these studies.

Lee Sproull earned her Ph.D degree at Stanford University; she is Associate Professor of Social Sciences at Carnegie-Mellon University. Her current research interests include organizational information processing and the social aspects of computing.

TABLE A.1.
Data Collection Characteristics for 1983 *Academy of Management Journal* Articles that Used Interviews or Questionnaires

Means of Data Collection	Number of Articles	Target Sample Size			Usable Returns			Usable Response Rates		
		Mean	s.d.	Range	Mean	s.d.	Range	Mean	s.d.	Range
Interview	1	n.g.	—	—	15	—	—	n.g.	—	—
Interview plus workplace questionnaire	5	151^a	93	86-217	153^b	66	80-208	95^a	2	93-96
Workplace questionnaire	12	335^c	402	42-1130	292^d	259	42-805	85^e	13	65-100
Mailed questionnaire	14	1032	1445	126-4550	477	656	58-2317	53	19	18-93
Workplace questionnaire plus mailed questionnaire	1	342	—	—	327	—	—	96%	—	—
Telephone interview	1	1023	—	—	622	—	—	61%	—	—

$^a n=2$
$^b n=3$
$^c n=6$
$^d n=11$
$^e n=5$

© Academy of Management Journal
1986, Vol. 29, No. 1, 170–182.

ATTITUDES TOWARD WORK SCHEDULES: CONSTRUCT DEFINITION, INSTRUMENT DEVELOPMENT, AND VALIDATION

RANDALL B. DUNHAM
University of Wisconsin
JON L. PIERCE
University of Minnesota-Duluth

Interest in the specific effects of work schedules has increased during recent years, focusing both on the traditional five-day/40-hour week (5/40) and on alternative work scheduling arrangements such as the four-day/40-hour week (4/40), the 4 1/2-day/40-hour week (4.5/40), staggered hours, and flextime. Practitioners and researchers have also examined various configurations of shiftwork schedules and have increased their efforts to document systematically the strengths and weaknesses of these schedules. Despite these efforts to understand more fully the effects of various schedules on workers and their organizations, most of the literature on alternative work schedules still focuses primarily on a relatively narrow range of workers' reactions to work schedules. Factors that have been studied include job satisfaction, stress, internal work motivation, job involvement, and organizational commitment (Glueck, 1979; Golembiewski & Proehl, 1978; Newstrom & Pierce, 1979; Pierce & Newstrom, 1982).

With few exceptions (cf. Allen & Hawes, 1979; Goodale & Aagaard, 1975; Mahoney, 1978; Nord & Costigan, 1973), practitioners and researchers have directed very little conceptual or empirical effort toward examining employees' attitudes regarding specific work schedule alternatives. This, in turn, has somewhat limited attempts to understand workers' attitudes toward alternative work schedules and how these attitudes may be used to design successful work schedule programs.

A 1975 study (Mahoney, Newman, & Frost) developed a set of scales to assess workers' attitudes toward a four-day work week. The results of this

---

The authors would like to thank the following for their contributions to this research: Alison Barber of General Mills, Lauri Breininger of Whirlpool, Michael L. Compton of Standard Oil, Rebecca A. Ellis of the University of Iowa, Kathy Chiaravalli of Data General, Jo Lynn Lund of James River Corporation-Dixie Northern Inc., Shelley C. Verbin of Pacific Gas and Electric, Ted Weinberger of the University of Minnesota-Twin Cities, and Randi K. Huntsman of Management Innovations. In addition, although they prefer to remain anonymous, we wish to thank the individuals from the local union and management of the midwestern utility company involved in this study for their extensive and enthusiastic support. Grants from the Department of Labor, the National Council on Employment Policy, and the University of Wisconsin Graduate School supported portions of this research.

study suggested promise for the assessment of such attitudes. Unfortunately, Mahoney and his colleagues did not show the items from their scales but only described them generally; nor did they provide reliability or validity information for the scales. To the best of our knowledge, these scales have not received additional development.

A more thorough study (Golembiewski, Yeager, & Hilles, 1975) developed a set of 18 items for assessing attitudes toward current work schedules. Although these items did not appear in their entirety, they apparently could be used with a wide range of types of current schedules. In fact, Golembiewski and coauthors used the set of items for a group of workers in reference to a 5/40 schedule and later in reference to a flextime schedule. This study also demonstrated the potential feasibility of systematic assessment of attitudes toward work schedules, although it had several limitations. First, the scales appeared to be designed to assess attitudes toward current work schedules, which disallowed the possibility of anticipating reactions to alternative work schedules until workers actually experienced them. Second, this paper did not describe the development of the items, so it is difficult to determine which domain the 18 items represented: the researchers' perspective, the subjects' perspectives, or some other perspective. Finally, determination of the factor structure for the 18 items was based on only 50 subjects—a marginal subject-to-item ratio at best. Thus, although this study demonstrated the importance and feasibility of assessing workers' attitudes toward work schedules, further work was clearly needed. To the best of our knowledge, to date no one has further developed these scales.

If an appropriate set of valid, reliable instruments were available to assess attitudes toward specific work schedules, organizations would be able to evaluate the strengths and weaknesses of their existing work schedules and to compare them with the anticipated strengths and weaknesses of various alternatives. Such instruments could also be used to select, design, implement, and evaluate alternative work schedules.

The primary purpose of our paper is to present a set of scales that can assess workers' attitudes toward almost any existing or alternative work schedule. We developed these scales during a three-phase research program involving six samples and over 700 clerical, technical, construction, manufacturing, and health care workers working under a variety of schedules. A related purpose of the paper is to explore the predictive effectiveness of several sets of variables for anticipating workers' preferences of work schedules. Finally, we hope that the general design used here can serve as a model for researchers and practitioners pursuing instrument development in other behavioral areas.

METHODS

Study Design

In phase 1 of the three-phase research program, the initial instruments were developed, administered, and tested for reliability and validity. The instruments were refined in phases 2 and 3, and further reliability and validity tests were conducted with additional samples.

Participants

Phase 1 employed three diverse samples—corporate office workers, traveling construction crews, and local repair crews and office workers—from a large midwestern utility company. These samples represented the population of workers for each of these three groups. Of a total 467 participants, approximately 23 percent were supervisory and 77 percent were nonsupervisory employees. The individuals participating in phase 2—office workers from the same utility company— were randomly selected by department. This sample consisted of 28 supervisors and 104 nonsupervisors. At the time of the study, all participants in both phases were working under a fixed 5/40 schedule. Phase 3 used samples from two county health departments at two times. The 156 nonsupervisory employees representing these two populations included nurses, educators, sanitarians, clerical workers, and administrators. One of these samples ($n = 57$) worked a flextime schedule during the entire study period. The second sample ($n = 99$) worked a fixed 5/40 schedule at time 1 and a 4/40 schedule at time 2.

Instrument Development

On the basis of literature on work schedules and interviews with a wide cross section of 60 workers from the organization involved in phase 1, we constructed an initial set of 100 attitude-based items. Together with a 12-member labor-management committee from the utility company sample, we reviewed and revised the list to obtain 66 items. A priori classification identified 18 items that focused on attitudes toward current schedules, 10 items that dealt with attitudes toward alternative schedules, and 38 items that related to perceived interference of work schedules with personal activities.

Instrument development in phase 2 built upon the findings from phase 1. Two new items were written and other items rewritten so that the same set of questions could address either existing or alternative schedules. The results from phase 2 suggested a need for a more reliable assessment of the effects of work schedules on outside constituents such as customers or clients than we had. Consequently, we wrote three new items for phase 3 to try to gain this information.

Data Collection

Groups of employees, who received work release time to participate in the study, completed paper-and-pencil questionnaires. Subjects participating in phase 1 answered questions regarding their attitudes toward their current 5/40 work schedule and about the level of interference with personal activities this schedule caused. In addition, each phase 1 group received the set of alternate work schedule items: the corporate office workers were asked to react to 4/40, 4.5/40, staggered, and flextime schedules;[1] traveling construction crews reacted to 4/40 and staggered schedules; local crews and office workers reacted to 4/40, 4.5/40, staggered, and flextime schedules.

In addition to the items related to work schedules, phase 1 participants answered questions relating to age, gender, educational level, marital status, number of children at home, personal security, company tenure, job involvement (Lodahl & Kejner, 1965), internal work motivation (Hackman & Oldham, 1975), job satisfaction[2] (Dunham, Smith, & Blackburn, 1977; Smith, 1976), discretion concerning work schedules (Pierce & Newstrom, 1982), and a direct measure of preference of work schedules.

Phase 2 participants answered the items dealing with interference with personal activities and the revised items on attitudes toward work schedules for both their existing 5/40 schedule and for a flextime work scheduling arrangement. Both samples of phase 3 participants completed the interference items and the second revision of the items on attitudes toward work schedules for both a fixed 5/40 and a 4/40 schedule. Two months later, we readministered the questionnaire to both phase 3 samples.

Analyses

Across the three phases of the research program, various analytic techniques examined the following characteristics of the instruments being developed: (1) dimensionality, (2) similarity of dimensionality across samples and types of schedules, (3) internal consistency, (4) test-retest reliability, and (5) ability to predict actual work schedule preferences. Because of differences in subject-to-item ratios, appropriateness of specific alternative schedules for particular groups of workers, and stages of scale development, we did not always use identical analytic procedures for all samples in each phase of the research program.

To identify scale dimensionality at phase 1, we used principal component analyses with varimax rotations to try to find relatively pure subdimensions of the general construct. In phases 2 and 3 oblimax rotations further refined the several moderately correlated dimensions. At all three phases, we established underlying numbers of factors through combined use of scree tests and interpretability criteria. We calculated congruency coefficients (Harmon, 1975) to compare solutions across samples, types of schedules, and time.

---

[1] The questionnaire described flextime as follows:

> There are two primary features to flextime plans:
>
> (1) There are one or more blocks of time (core hours) each day during which all employees must be at work. For example, all employees might be required to be at work between 10:00 A.M. and noon and between 2:00 P.M. and 3:00 P.M.
>
> (2) On a daily basis, each worker can decide when to work the remaining hours of the work day. Thus, each employee is free to choose his/her starting time, break time, and quitting time as long as he/she works eight hours a day and is present during the core hours.

[2] The short form of the Index of Organizational Reactions was used.

On the basis of the results of the component analyses, unit-weighted scale scores were constructed. We calculated coefficient alpha estimates (Cronbach, 1951) to assess the internal consistency for each scale and examined test-retest reliability estimates for the phase 3 samples.

As a partial test of the construct validity of the new scales, a nomological network was created including the constructs as measured by the new scales, actual work schedule preferences, and several other potential predictors of schedule preferences. Predictors included personal characteristics, job involvement, work motivation, satisfaction, and current discretion over work scheduling. With data from the phase 1 corporate office workers sample, we used a series of discriminant function analyses with classification analyses and multivariate omega-squared estimates (Tatsuoka, 1973) to approximate the nature, strength, and predictive accuracy of each set of variables in the network for the prediction of actual work schedule preferences.

RESULTS

Attitudes Toward Work Schedules

Phase 1. The 18 items assessing attitudes toward the current work schedule and the 10 items directed toward each of the alternate schedules were subjected to component analyses with varimax rotations for the combined sample of 467 and for the two individual samples with sufficient subject-to-item ratios. Each analysis of attitudes toward current work schedules produced a four-factor solution. Factor congruency estimates used to compare the similarity of the solutions across the samples ranged from .79 to .93 for the first three factors and between .39 and .73 for the fourth. Fifteen of the 18 items subsequently defined four unit-weighted scales designated as personal effects, organization effects, coordination of work effects, and perceived uniqueness of work schedule. Coefficient alpha estimates ranged from .74 to .95 for the first three scales and from .59 to .69 for the fourth scale.

Each analysis of the ten items on attitudes toward alternative work schedules suggested a two-factor solution. Factor congruency estimates for the solutions across samples and types of schedules ranged from .96 to .99 for the first factor and from .97 to .99 for the second. We used two of these items to define a unit-weighted scale interpreted as work coordination effects; four items defined a scale interpreted as personal effects; and four items failed to play meaningful roles because of low or complex loading. Coefficient alpha estimates for these two dimensions ranged from .80 to .96.

The results of these analyses identified several shortcomings that caused us to revise the instruments, with help from the labor-management committee. Of particular concern were (1) low congruency and reliability estimates for the fourth current schedule scale; (2) the ability to identify only two dimensions for alternative schedules; and (3) a recognition that it would be very advantageous to have a single set of items for assessing attitudes toward both current or alternative work schedules.

Phase 2. As a result of the work done in phase 1, 20 revised items emerged as usable with any current or alternative schedule. In phase 2, these items were used to obtain reactions to both 5/40 and flextime schedules. Component analyses with oblimax rotations suggested very similar five-factor solutions for both analyses. Unit-weighted scale scores were formed for these dimensions, identified as: general affect toward schedule, work coordination effects, uniqueness of work schedule, family attitude toward schedule, and schedule's effects on social and family life. Coefficient alpha estimates were very good for the first three scales, ranging from .77 through .93. The fourth scale had an alpha of .78 for the 5/40 scale, but only .60 for the flextime scale. The fifth scale had low alphas (.40s) for both schedules.

Two circumstances—the reliability problems identified in phase 2 and the inability of the instrument to identify adequately a dimension dealing with the effects of a schedule on outside constituents—led to the second major revision of these items and the research conducted in phase 3. Phase 3 was also intended to extend assessment of the external validity and stability of the research scales.

Phase 3. The second major revision of the scales on attitudes toward schedules resulted in a 23-item instrument. Two of the items in this set appeared inappropriate for respondents who were not married or who were living alone. These items stated "In general my family would be satisfied with my working..." and "My husband/wife would like me to work...." As a result, we decided to exclude these two items from analyses and from future versions of the scales.

Component analyses with oblimax rotations for each of the phase 3 combinations of sample, schedule, and time suggested that a six-factor solution was most appropriate. Of the six factors (see Table 1), four were parallels of factors from phase 2. These were general affect toward schedule, work coordination effects, uniqueness of schedule, and schedule's effects on social and family life. In addition, two factors not defined in phase 2 were identified: effects on transportation and personal security, and effects on service to external constituencies.[3]

Unit-weighted scales for the six dimensions across samples and types of schedules produced coefficient alphas that were for the most part at clearly acceptable levels (.70s and .80s). The major exception was the transportation and personal security scale, with alphas ranging from .29 to .90. Test-retest coefficients were also mostly strong (.60s to .80s) with a few exceptions. In particular, coefficients for schedule's effects on social and family life were poor, averaging in the .20s.

Perceived Interference

The 38 items assessing perceived interference with personal activities were analyzed for the combined sample from phase 1. The results of the

---

[3] The term external constituencies is somewhat awkward, but we use it here to deal with service to those outside of home organizations. This might include customers, clients, suppliers, regulatory agency representatives, and so forth.

TABLE 1
Results of Principal Component Analysis[a] of Final Instrument Measuring Attitudes toward Work Schedules

	Factors					
Items[b]	**1**	**2**	**3**	**4**	**5**	**6**
General affect toward schedule						
1. I would be very productive under the _________ method of scheduling.	**89**	05	-05	03	14	04
2. The _________ schedule would have a favorable influence on my overall attitude toward my job.	**80**	00	-08	-10	01	08
3. My current work schedule would encourage me to do my best.	**78**	01	-09	05	-12	-06
4. Taking everything into consideration, I would be satisfied with my life in general while working a _________ work week.	**68**	09	-02	-25	-04	07
5. I would be dissatisfied with my current work schedule.	**-68**	00	04	14	-01	03
6. I personally would like the _________ method of scheduling.	**62**	12	-08	-15	-01	-03
7. A _________ schedule would have an unfavorable influence on my physical health.	**54**	23	22	09	-20	-14
Service to external constituency						
8. The _________ schedule would improve client access to the services of the Department.	05	**85**	-01	00	00	05
9. The _________ method of scheduling would help the Department meet the needs of its clients.	-08	**82**	-10	-01	-04	-08
10. The _________ method of scheduling would make it easier for me to meet the service needs of the clients.	10	**76**	-01	-12	01	-02
11. The _________ method of scheduling would hurt the quality of client service.	-09	**-53**	43	-07	-10	04
Work coordination effects						
12. The _________ method of scheduling would cause problems in coordinating work with my supervisor.	02	-04	**83**	07	11	-05
13. The _________ method of scheduling would cause problems in coordinating work with my co-workers.	-11	-12	**81**	03	04	05

TABLE 1 (continued)

	Factors					
Items[b]	1	2	3	4	5	6
Schedule effects on social and family life						
14. The _________ work schedule would have an unfavorable influence on my family life.	08	-03	08	**86**	05	04
15. The _________ method of scheduling would make it easier for me to coordinate my schedule with the schedules of other family members.	14	09	08	**-66**	07	-05
16. The _________ work schedule would have an unfavorable influence on my social life.	-38	21	19	**51**	10	11
Uniqueness of work schedule						
17. Almost none of my neighbors work a _________ work schedule.	-08	13	03	-10	**85**	12
18. Almost all of my friends work a _________ work schedule.	-11	14	-07	-13	**-78**	20
Transportation and personal security						
19. Working or leaving the work place at an unusual (early or late) hour would represent a personal security problem for me under the _________ work schedule.	18	-09	-10	09	-03	**89**
20. Under the _______ method of scheduling, transportation to and from work would be difficult.	-31	10	25	01	-03	**57**
Item not used in scales						
21. Overall, the _________ method of scheduling would have more good than bad aspects.	37	53	11	-08	-04	02

[a] Decimal points omitted.

[b] A 5-point response format was used, ranging from 1 = strongly disagree to 5 = strongly agree.

principal component analysis with varimax rotations suggested three underlying dimensions (see Table 2). Twenty-one of the items defined three scales: interference with activities involving family and friends; interference with access to services, events, and consumer goods; and interference with financial and related activities. With one exception, coefficient alpha estimates for the three scales for the three samples and for the combined sample were very high (.81 to .93).

TABLE 2
Results of Principal Component Analysis: Interference with Personal Activities[a]

Items[b]	Factors 1	2	3
Interference with activities involving family and friends			
1. Spend time with your children.	**84**	33	14
2. Maintain personal family relations.	**76**	30	22
3. Spend time with your spouse.	**76**	29	24
4. Spend time with friends.	**75**	35	17
5. Take family on recreational outings.	**75**	33	21
6. Take part in children's education.	**73**	20	31
7. Attend children's school activities.	**68**	21	35
8. Participate in sports with others.	**66**	33	27
9. Participate in non-work organizations.	**64**	31	24
Interference with access to services, events, consumables			
10. See a movie.	25	**78**	16
11. Go to tavern or buy liquor.	19	**75**	07
12. Go to the laundromat.	31	**70**	24
13. Go out to eat.	21	**68**	31
14. Attend play, concert, cultural event.	41	**68**	19
15. Watch favorite TV programs.	33	**61**	09
16. Buy groceries.	31	**58**	33
17. Go shopping for clothes.	37	**53**	39
18. Use public transportation.	23	**45**	21
Interference with activities involving family and friends			
19. Go to the bank.	18	14	**79**
20. Cash a check.	10	18	**75**
21. Go to the post office.	30	07	**65**
Items not used in scales			
22. Attend spectator sports.	43	47	29
23. Get haircut or styled.	38	33	43
24. Pay bills.	24	49	53
25. Get prescription filled.	23	47	63
26. Go to the library.	21	41	54
27. Visit doctor or dentist.	38	14	62
28. Buy gas.	33	45	38
29. Continue your education.	55	30	18
30. Have repair work done.	57	09	43
31. Visit relatives.	63	39	11
32. Do individual sports or hobbies.	68	34	31
33. Attend family events.	56	35	17
34. Go to the dry cleaners.	37	32	49
35. Attend church or synagogue.	41	58	10
36. Do routine work around home.	68	24	37
37. Do part-time military activity.	46	38	25
38. Spend time alone.	37	35	19

[a] Decimal points omitted.

[b] Some items are abbreviated in this table; the complete items are available upon request. Items were preceded by "How easy or difficult is it for you now to: . . . ." Response format was from 1 = very difficult to 5 = very easy.

Prediction of Work Schedule Preference

In the corporate office worker sample, several sets of potential predictors of work schedule preferences were used to predict workers' actual first choices from a set of five work schedules that were under consideration. The sets of predictors included: personal characteristics, job involvement and motivation, job satisfaction, current discretion over work schedules, interference with personal activities, and attitudes toward schedules, as measured with the newly developed instruments. A series of discriminant function analyses determined whether a given set had a significant relationship with schedule preference, identified the percent of variance explained, and assessed classification accuracy using the set of variables to predict schedule preference.

Neither the job satisfaction nor the job involvement/motivation sets of variables had significant relationships with schedule preferences. Personal characteristics accounted for 3 percent of the variance with a 39 percent classification accuracy. Current discretion over the work schedule also accounted for 3 percent of the variance with a 30 percent classification accuracy. Interference from a work schedule explained 12 percent of the variance and had a 23 percent classification accuracy. The set of attitudes toward schedules explained 74 percent of the variance and produced a 67 percent accuracy in classification. The combined set of variables accounted for 85 percent of the variance and provided correct classification in 71 percent of all cases.

DISCUSSION

This paper reports the development and validation of a set of scales to measure workers' attitudes toward various work schedules. This three-phase investigation resulted in the identification of scales to measure six attitudes toward work schedules: general affect toward a schedule; work coordination effects; uniqueness of schedule; schedule's effects on social and family life; effects on transportation and personal security; and effects on service to external constituencies. In addition, we developed measures to assess perceived interference with three sets of personal activities: family and friends; access to services, events, and consumables; and financial and related activities. We found it particularly interesting that, in forming attitudes toward work schedules, workers focused not only on personal and family effects but also on the consequences that particular schedules were likely to have on organizational effectiveness in terms of both work coordination and service to external constituencies. Interference caused by work schedules can be in one or more of the three distinct areas. We distinguished among these three different areas of interference because managerial actions taken to remedy associated problems may be very different, depending on where interference is most severe.

The scales developed to measure each of the preceding dimensions have been shown to have adequate validity and reliability. With the possible exception of the uniqueness scale, the measure warranting the most serious remaining questions about reliability, we believe that the scales developed here can be used with a wide range of employee types or schedules and can maintain adequate reliability and validity. Since the same scales can apply to many combinations of existing and alternative schedules, researchers can make direct comparisons for each of the attitudes across types of schedules.

Placement of the attitude scales into a nomological network provided additional construct validity evidence. We showed that wide differences in preferences toward work schedules could not be explained entirely on the basis of employees' characteristics or general affective orientations, as the nonempirical literature has often suggested. Attitudes must actually be measured; it is not adequate merely to draw inferences based on employee characteristics. Once schedule attitudes are assessed, we can predict work schedule preferences with high degrees of accuracy. Classification analyses conducted in these studies were able to predict work schedule preferences with an overall accuracy of up to 71 percent. The strongest predictors of work schedule preferences were specific attitudes toward work schedules—primarily affect—followed by personal characteristics, current scheduling discretion, and interference with personal activities. In many ways, it is these findings that most strongly support the construct validity of the measures developed in these studies.

An important remaining test of the utility of the scales developed here will involve using these scales to plan, implement, and evaluate alternative work schedules. This is being done in a series of studies currently nearing completion (Dunham, Pierce, & Castañeda, 1985). In one of these studies, the scales identified a flextime schedule as preferable to the existing fixed 5/40 schedule. Using a quasi-experimental design, researchers used the scales prior to a work schedule change for both an experimental and control group and used them again three and six months after the introduction of flextime. The scales isolated a number of significant improvements in attitudes following the introduction of flextime. Perhaps more important, in terms of the validity of the scales, the pretest measures of flextime attitudes correctly anticipated subsequent reactions after flextime was actually introduced. In a second study to be reported by Dunham and colleagues (1985), the scales were used with a natural quasi-experimental design for a group of workers who changed from a fixed 5/40 to a 4/40 and back to a 5/40 schedule. Again the scales isolated a number of significant attitudinal changes, and again the investigators correctly anticipated these reactions prior to the introduction of schedule changes.

Practitioners should find the instruments reported here of use in designing or modifying work schedules. More complete knowledge of workers' complex attitudes toward work schedules will allow a much more systematic approach to work schedule design and redesign than has been possible in the past. The fact that workers' attitudes toward current and alternative work schedules focus both on the personal and organizational effects of these schedules suggests that workers' attitudes and opinions can be very useful in designing work schedules that benefit organizations as well as workers. These scales can help organizations identify desirable alternative work schedules, anticipate probable positive and negative aspects of schedules, and systematically evaluate the effectiveness of alternative work schedule interventions.

REFERENCES

Allen, R. E., & Hawes, D. K. 1979. Attitudes toward work, leisure, and the four-day workweek. ***Human Resources Management***, 18 (1): 5–10.

Cronbach, L. J. 1951. Coefficient alpha and the internal structure of tests. ***Psychometrika***, 16: 297–334.

Dunham, R. B., Pierce, J. L., & Castañeda, M. 1985. ***Alternative work schedules: Two field experiments***. Working paper, Graduate School of Business, University of Wisconsin, Madison.

Dunham, R. B., Smith, F. J., & Blackburn, R. S. 1977. Validation of the Index of Organizational Reactions with the JDI, the MSQ, and FACES scales. ***Academy of Management Journal***, 20: 420–432.

Glueck, W. F. 1979. Changing hours of work: A review and analysis of the research. ***Personnel Administrator***, 24 (3): 44–67.

Golembiewski, R. T., & Proehl, C. W., Jr. 1978. A survey of the empirical literature on flexible workhours: Character and consequences of a major innovation. ***Academy of Management Review***, 3: 837–853.

Golembiewski, R. T., Yeager, S., & Hilles, R. 1975. Factor analysis of some flexitime effects: Attitudinal and behavioral consequences of a structural intervention. ***Academy of Management Journal***, 18: 500–509.

Goodale, J. G., & Aagaard, A. K. 1975. Factors relating to varying reactions to the 4-day workweek. ***Journal of Applied Psychology***, 60: 33–38.

Hackman, J. R., & Oldham, G. R. 1975. Development of the Job Diagnostic Survey. ***Journal of Applied Psychology***, 60: 159–170.

Harmon, H. H. 1975. ***Modern factor analysis*** (3rd ed.). Chicago: University of Chicago Press.

Lodahl, T. M., & Kejner, M. 1965. The definition and measurement of job involvement. ***Journal of Applied Psychology***, 49: 24–33.

Mahoney, T. A. 1978. The rearranged work week: Evaluation of different work schedules. ***California Management Review***, 20 (4): 31–39.

Mahoney, T. A., Newman, J. M., & Frost, P. J. 1975. Worker perceptions of the four-day week. ***California Management Review***, 18 (1): 31–35.

Newstrom, J. W., & Pierce, J. L. 1979. Alternative work schedules: The state of the art. ***Personnel Administrator***, 24 (10): 19–23.

Nord, W. R., & Costigan, R. 1973. Work adjustment to the four-day week: A longitudinal study. ***Journal of Applied Psychology***, 58: 60–66.

Pierce, J. L., & Newstrom, J. W. 1982. Employee responses to flexible work schedules: An inter-organization, inter-system comparison. ***Journal of Management***, 8 (1): 9–25.

Smith, F. J. 1976. Index of organizational reactions (IOR). ***JSAS Catalog of Selected Documents in Psychology***, 6: No. 1265.

Tatsuoka, M. M. 1973. ***An examination of the statistical properties of a multivariate measure of strength of relationship***. Urbana: University of Illinois.

Randall B. Dunham earned his Ph.D. degree at the University of Illinois; he is currently a professor at the University of Wisconsin-Madison in organizational behavior and industrial relations. His current research interests include work scheduling, job design, organizational commitment, and organizational change.

Jon L. Pierce earned his Ph.D. degree at the University of Wisconsin-Madison; he is currently a professor of organization and management at the University of Minnesota-Duluth. His current research interests include work scheduling, job design, organizational commitment, and organizational change.

Please send all address corrections to:
WALTER B. NEWSOM
Secretary - Treasurer
ACADEMY OF MANAGEMENT
P.O. DRAWER KZ
Mississippi State, MS 39762

© *Academy of Management Journal*
1986, Vol. 29, No. 1, 183–188.

STYLE GUIDE FOR AUTHORS

ARTICLES

Submit four copies of the manuscript; be sure that they are good, clear copies and that all pages are included in each copy. The manuscript should be typed on standard size (8½" x 11") paper, *double-spaced throughout* (including footnotes, references, quotations, and appendices), on only one side of the paper, and with wide margins (one inch or more) at top, bottom, and both sides of each page. Manuscripts prepared on computers should be printed on letter quality printers or, if other printers are used, in double-strike or enhanced print. Footnotes, references, appendices, tables, and figures should be on separate sheets of paper and should be arranged at the end of the manuscript in the order listed in this sentence. There is no absolute limit, but the length of articles should not ordinarily exceed 30 manuscript pages, including references, appendices, tables, and figures.

Title Page and Abstract

The first page of the manuscript should include the title of the article (typed in all capital letters), the authors' names (typed in all capitals), their affiliations, addresses, and phone numbers (typed with initial caps only). Example:

THE USE OF COMPUTERS IN MANAGERIAL DECISION MAKING

JANICE M. BEYER
School of Management
State University of New York at Buffalo
Jacobs Management Center
Buffalo, New York 14260
(716 636-3229)

No mention of authors' names should be made in the body of the paper, except where appropriate in citations and the references.

The second page, numbered page 2, should repeat the title and include a 75-word abstract of the article. The text of the article should begin on page 3. Page numbering should continue through all pages of the manuscript, including those with footnotes, references, appendices, tables, and figures.

Acknowledgments

An unnumbered footnote that will be printed at the bottom of the first page can be used to acknowledge financial support and/or assistance of others in preparing the manuscript. In the manuscript, the text for this footnote should appear at the bottom of the same page as the abstract (page 2). Like other footnotes, it should be separated from the abstract by a 10-space line beginning at the left-hand margin.

Footnotes

Other footnotes should be used sparingly. Minimize their use for parenthetical discussion; material that is pertinent can often be integrated into the text. They should *not* be used for citing references (see References below). The text for all footnotes should appear on separate pages at the end of the body of the article.

Headings

Main headings should be used to designate the major sections of the article; three or four main headings should be sufficient for most articles. Initial headings, such as "Introduction," are unnecessary. Main headings should be centered on the page and typed in all capitals. They should *not* be underlined. Example:

RESEARCH DESIGN

Secondary headings should be typed flush with the left margin and in small letters, with major words beginning with capitals. Secondary headings should *not* be underlined. Example:

Sampling Procedures

Third-order or paragraph headings should begin with a standard paragraph indention and be typed in capital and small letters, with only the initial word capitalized. Paragraph headings should be followed by a period; they should *not* be underlined. Example:

Manager sample. Respondents consisted of a random sample of 300 managers . . .

The text should follow on the same line.

Tables

Tables should be typed, double-spaced, on separate pages (one page for each table) from the text. They should be grouped together following the references. Each table should have the word TABLE (typed in all caps) and its number (arabic numerals) centered at the top. The table title should be in capital and small letters and centered on the page directly under the table number. Example:

TABLE 1
Firms in Sample

Tables should be numbered consecutively from the beginning to the end of the article. The position of the table in the manuscript should be indicated in the text as follows:

Insert Table 1 about here

Footnotes to tables are of two types:

(1) General footnotes that explain the table as a whole, the designations of table columns or rows, or an individual item. All of these should be designated by superscript small letters (^{a,b,c}), with the footnotes for each separate table beginning with ^{a}.

(2) Footnotes used to indicate the level of significance. These should follow any other footnotes and be designated by one or more asterisks. One asterisk should be used when only one level of probability is indicated. When more than one level is indicated, a single asterisk should be used for the level of lesser significance (* for $p < .05$ and ** for $p < .01$).

If it is necessary to distinguish some numerals in a table from others (for example, to indicate which factor loadings define a factor), boldface type can be used. In the typed manuscript, any numerals that should be set in boldface type should be underlined with a wavy line. This possibility should *not* be used when other conventions, such as footnotes, are sufficient.

Figures

Because graphs, charts, and other figures may be expensive to reproduce in the *Journal*, they should be used only where they contribute substantially to an understanding of the research presented beyond what is possible in the textual presentation.

The spacing and lettering used in figures should allow for subsequent reduction in size by as much as 50 percent so that the figure will fit the size of the *Journal's* page. The original artwork for figures should not be submitted until after the manuscript has been accepted for publication.

Figures should be numbered consecutively with arabic numerals and their position in the text indicated as for tables (see above). Each figure should be presented on a separate page with FIGURE (typed in all caps) and its number centered above it and a short identifying title in capital and small letters centered underneath the figure number. Example:

FIGURE 1
Model of Contextual Influences on Performance

References

An alphabetically ordered list of references cited in the text should be included at the end of the article. References should begin on a separate page headed REFERENCES. Continue the pagination.

Entries in the *list of references* should be alphabetized by last name of author (first author if more than one) or editor, or by the corporate author (U.S. Census Bureau) or periodical name (*Wall Street Journal*) if there is no indication of individual authors or editors. Several references by an identical author (or group of authors) are ordered by year of publication, with the earliest listed first. Authors' names are repeated for each entry.

Citations to references should be designated throughout the text by enclosing the authors' names and the year of the reference in parentheses. Example:

> Several studies (Adams, 1974; Brown & Hales, 1975, 1980; Collins, 1976) support this conclusion.

(Note alphabetical order and use of ampersand.) Page numbers must be included in the citation to provide the exact source of a direct quotation. Example:

> Adams has said that writing a book is "a long and arduous task" (1974: 3).

They should also be used when specific arguments or findings of authors are paraphrased or summarized. As indicated in the second example, if the name of the author occurs in the body of the sentence, only the year of publication is cited in parentheses. Otherwise, as indicated in the first example, both name and date appear in parentheses, separated by a comma. Page numbers follow the date of publication given in parentheses and are separated from it by a colon.

If a work has two authors, always cite both names every time the work is cited in the text. If the work has more than two authors, cite all authors the first time the reference occurs; in subsequent citations of the same work, include only the surname of the first author followed by "et al." (not underlined) and the year. Examples:

> Few field studies use random assignment (Franz, Johnson, & Schmidt, 1976).
> (first occurrence)
>
> . . . even when random assignment is not possible (Franz et al., 1976: 23).
> (subsequent citation)

However, for works with six or more authors, use only the surname of the first author followed by et al. whenever the work is cited.

Book entries in the list of references follow this form: Authors' or Editors' Last Names, Initials. Year. Title of book. (Book titles are underlined and typed in lower case letters except for the first word and the first word after a colon.) City Where Published, State or Country (only if necessary to identify the city): Name of Publisher. Examples:

Boulding, K.E. 1956. The image. An Arbor: University of Michigan Press.

Kahn, R.L., & Boulding, E. (Eds.) 1964. Power and conflict in organizations. Glencoe, Ill.: Free Press.

Katz, D., & Kahn, R. L. 1978. The social psychology of organizations (2nd ed.). New York: Wiley.

U.S. Department of Labor Statistics. 1976-1983. Employment and earnings. Washington, D.C.: U.S. Government Printing Office.

Periodical entries follow this form: Author's Last Name, Initials. Year. Title of article or paper (in lower case letters except for the first word and the first word after a colon). Name of Periodical, volume number (issue number): page numbers. Examples:

Fry, L.W., & Slocum, J.W., Jr. 1984. Technology, structure, and workgroup effectiveness: A test of a contingency model. Academy of Management Journal, 27: 221-246.

Goggin, W. C. 1974. How the multidimensional structure works at Dow Corning. Harvard Business Review, 55 (1): 54-65.

The issue number should only be included if the periodical pages are *not* numbered consecutively throughout the volume; i.e., if each issue begins with page 1.

If a periodical article has no author, the name of the periodical should be treated like a corporate author, both in the citation and in the references. For example:

There is fear that Social Security rates may rise (Wall Street Journal, 1984).

Wall Street Journal. 1984. Inflation rate may cause Social Security increase. September 24: 14.

Chapters in books follow this form: Author's Last Names, Initials. Year. Title of article (in lower case except for the first word and the first word after a colon). In Editors' Initials and Last Names (Eds.), Title of book: page numbers. City Where Published, State or Country (only if necessary to identify the city): Name of Publisher. Examples:

Berg, N. A. 1973. Corporate role in diversified companies. In B. Taylor & N. MacMillan (Eds.), Business policy: Teaching and research: 298-347. New York: Wiley.

Roberts, F. S. 1976. Strategy for the energy crisis: The case of commuter transportation policy. In R. Axelrod (Ed.), Structure of decision: 142-179. Princeton, N.J.: Princeton University Press.

Unpublished papers, dissertations, and presented papers should be listed in the references using the following formats:

Duncan, R.G. 1971. Multiple decision making structures in adapting to environmental uncertainty. Working paper No. 54-71, Northwestern University Graduate School of Management, Evanston, Ill.

Smith, M. H. 1980. A multidimensional approach to individual differences in empathy. Unpublished doctoral dissertation, University of Texas, Austin.

Wall, J. P. 1983. Work and nonwork correlates of the career plateau. Paper presented at the annual meeting of the Academy of Management, Dallas, Texas.

Appendices

Lengthy but essential methodological details, such as explanations of long lists of measurements, should be presented in one or more appendices at the end of the article. This material should be presented in as condensed a form as possible; full sentences are not necessary. No tables should be included in the appendices. A single appendix should be titled Appendix in all caps. If more than one appendix is needed, they should be titled and ordered alphabetically: Appendix A, Appendix B, etc.

Biographical Sketches

At the time an article is accepted for publication, a brief biographical sketch of 50 words or less should be submitted for each author. It should include where highest degree was earned, present position, affiliation, and current research interests. For example:

Andrea Barber earned her Ph.D. degree at the University of Wisconsin; she is an associate professor of management and Director of the Management Improvement Center at Famous University, Oxbridge, Ohio. Her current research interests include dual-career families and sociotechnical systems in organizations.

RESEARCH NOTES

Research notes contain brief descriptions of original research. To be considered for the Research Notes section, manuscripts should not exceed ten double-spaced typewritten pages in length. Descriptive surveys, replications, methodological demonstrations or analyses, studies that make incremental advances to established areas of inquiry, and commentaries with new empirical content are especially appropriate.

Manuscripts intended for the Research Notes section should be prepared according to the above instructions for articles, except that an abstract is not needed.

Harvard Business School Press

Our business is Business . . .

Managing in the Service Economy
James L. Heskett
The service industry–including banking, communications, law, transportation, fast-foods, and consulting–is the fastest growing sector in today's declining economy. Harvard Business School professor James L. Heskett presents the "strategic service vision"–a practical set of strategies that any manager can follow for success. Anecdotes and case studies of leading service companies as well as clear, workable instructions make *Managing in the Service Economy* an *invaluable* guide for today's service manager.
May 240 pp. Tables Figures Index
ISBN 0-87584-130-9 $19.95 Cloth

Investment Banking and Diligence
What Price Deregulation?
Joseph Auerbach and Samuel L. Hayes, III
A thorough and provocative analysis of the implications of recent regulations that allow new securities to be sold immediately after registration. Important reading for corporate managers, lawyers, investment bankers and "Wall Street watchers."
April 288 pp. Tables Index
ISBN 0-87584-171-6 $24.95 Cloth

Leadership and Ambiguity
Second Edition
Michael D. Cohen and James G. March
with a new preface by the authors
This classic study presents both a thorough and witty discussion of the role of the college president and a broadly applicable analysis of the problems facing all leaders of what the authors term "organized anarchies."
February 290 pp. Tables Index
ISBN 0-87584-174-0 $19.95 Cloth
ISBN 0-87584-131-7 $9.95 Paper

Introducing . . .
Harvard Business School Classics

High-quality paperback reprints bring groundbreaking research from Harvard Business School to contemporary managers. Each new edition includes a new preface by the authors commenting on the influence of their work on modern management techniques.

Directors: Myth and Reality
Myles L. Mace
An insider's view of the corporate boardroom–and why it doesn't work today.
May 224 pp. Index
ISBN 0-87584-125-2 $9.95 Paper

Strategy, Structure, and Economic Performance
Richard P. Rumelt
A correlation of how diversification affects the economic performance of a company.
May 235 pp. Tables Index
ISBN 0-87584-126-0 $9.95 Paper

Managing the Resource Allocation Process
Joseph L. Bower
An analysis, from a manager's point of view, of the factors that influence corporate allocation decisions.
May 382 pp. Tables Index
ISBN 0-87584-128-7 $12.95 Paper

Organization and Environment
Paul R. Lawrence and Jay W. Lorsch
Provides a key to understanding which organizational characteristics lead to success in different businesses.
May 296 pp. Tables/Figures Index
ISBN 0-87584-129-5 $12.95 Paper

Strategy for Financial Mobility
Gordon Donaldson
A classic study of how companies can improve their financial stability by preplanning for possible crises.
May 343 pp. Tables Index
ISBN 0-87584-127-9 $12.95 Paper

Harvard Business School Press • Boston, Massachusetts

Distributed by Harper & Row Publishers, Inc.
To order, call toll-free, 1-800-242-7737, or write: Harper & Row, Keystone Industrial Park, Scranton, PA 18512.

It's none of our business.

After all, you're the teacher.
No one knows better than you do what your students need,
and what text best fits their needs.

At West, we work hard to provide you
with a wide variety of texts to choose from.
Texts that adapt to the way you want your students to learn.
So you can do the teaching *your* way.

That's *our* business.

The 1986 management texts from West.

Introduction to Management

Management: Concepts and Effective Practice, 2e by Michael A. Hitt, Texas A&M University, R. Dennis Middlemist, Colorado State University, and Robert L. Mathis, University of Nebraska, Omaha.
Instructor's Manual and Test Bank/Transparency Masters/Study Guide/Videotape/MicroTest-WESTEST /West Study. 600 pp, hardcover.

Fundamentals of Production/ Operations Management, 4e by Harold E. Fearon, William A. Ruch, and Vincent G. Reuter, all of Arizona State University, C. David Wieters, New Mexico State University, and Ross R. Reck.
Instructor's Manual/Test Bank. 180 pp, softcover.

Organizational Behavior

Organizational Behavior, 4e by Don Hellriegel, Texas A&M University, John W. Slocum, Jr., Southern Methodist University, and Richard W. Woodman, Texas A&M University.
Instructor's Manual/Test Bank/Transparency Masters/Study Guide. 650 pp, hardcover.

Organizational Behavior: Experiences and Cases by Dorothy Hai, University of Wisconsin/LaCrosse.
To accompany Organizational Behavior, 4e. Instructor's Manual. 300 pp, softcover.

Strategic Management

The West Series in Strategic Management. Series includes:
Strategy Formulation: Power and Politics, 2e by Ian MacMillan, New York University, and Patricia Jones, Management Consultant.
Setting Strategic Goals and Objectives, 2e by Max D. Richards, Rollins College.
Strategy Implementation: Structure, Systems, and Process, 2e by Jay R. Galbraith, Management Consultants LTD, and Robert K. Kazanjian, University of Michigan.
Strategic Control by Peter Lorange, Wharton School, University of Pennsylvania, Michael F. Scott Morton, Sloan School, Massachusetts Institute of Technology, and Sumantra Goshal, INSEAD, Fountainbleau, France.

Macroenvironmental Analysis for Strategic Management by Liam Fahey, Boston University, and V.K. Narayanan, University of Kansas.
Strategy Formulation: Issues and Concepts by Charles W. Hofer, University of Georgia.
200 pp/book, softcover.

Personnel

Effective Personnel Management, 2e by Randall S. Schuler, New York University, and Stuart A. Youngblood, Texas A&M University.
Instructor's Manual and Test Bank/Study Guide/ Software and PC Manual/ Transparency Masters. 700 pp, hardcover.

Case Problems in Personnel and Human Resource Management by Randall S. Schuler and Stuart A. Youngblood.
Instructor's Manual. 325 pp, softcover.

Productivity Through People by William B. Werther, Jr., and William A. Ruch, both of Arizona State University, and Lynne McClure, McClure and Associates.
500 pp, softcover.

Organizational Theory

Organizational Theory and Design, 2e by Richard L. Daft, Texas A&M University.
Instructor's Manual/Test Bank. 571 pp, hardcover.

Marketing

Transportation, 2e by John Coyle, Pennsylvania State University, Edward Bardi, University of Toledo, and Joseph Cavinato, Pennsylvania State University.
Instructor's Manual and Test Bank/ Transparency Masters. 550 pp, hardcover.

Analysis for Strategic Market Decisions by George S. Day, University of Toronto.
Instructor's Manual. 270 pp, softcover.

For more information on these and other West texts,
or to order an examination copy, write to
West Publishing Company, College Division (KS),
P.O. Box 64526, St. Paul, MN 55164-1003,
or call toll-free 1-800-328-9424
(in Minnesota and Canada call 612-228-2778).

®
West Publishing Company.

NEW FROM ELSEVIER!

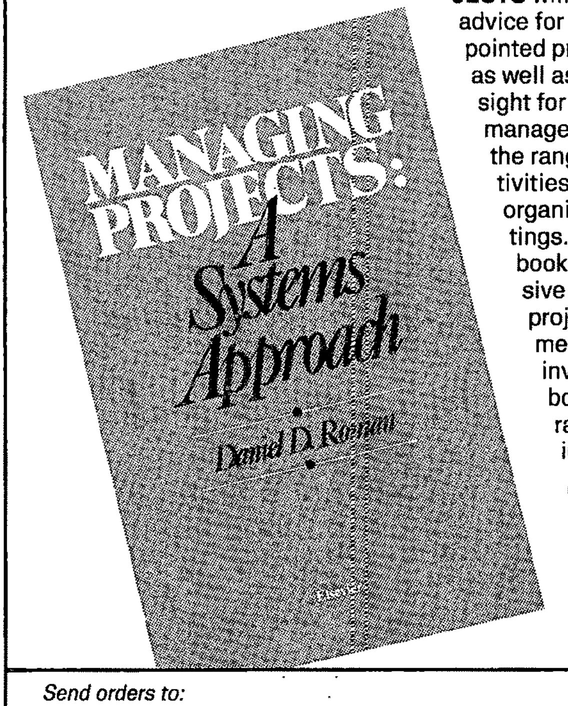
MANAGING PROJECTS:
A Systems Approach

Daniel D. Roman, Professor Emeritus, The George Washington University

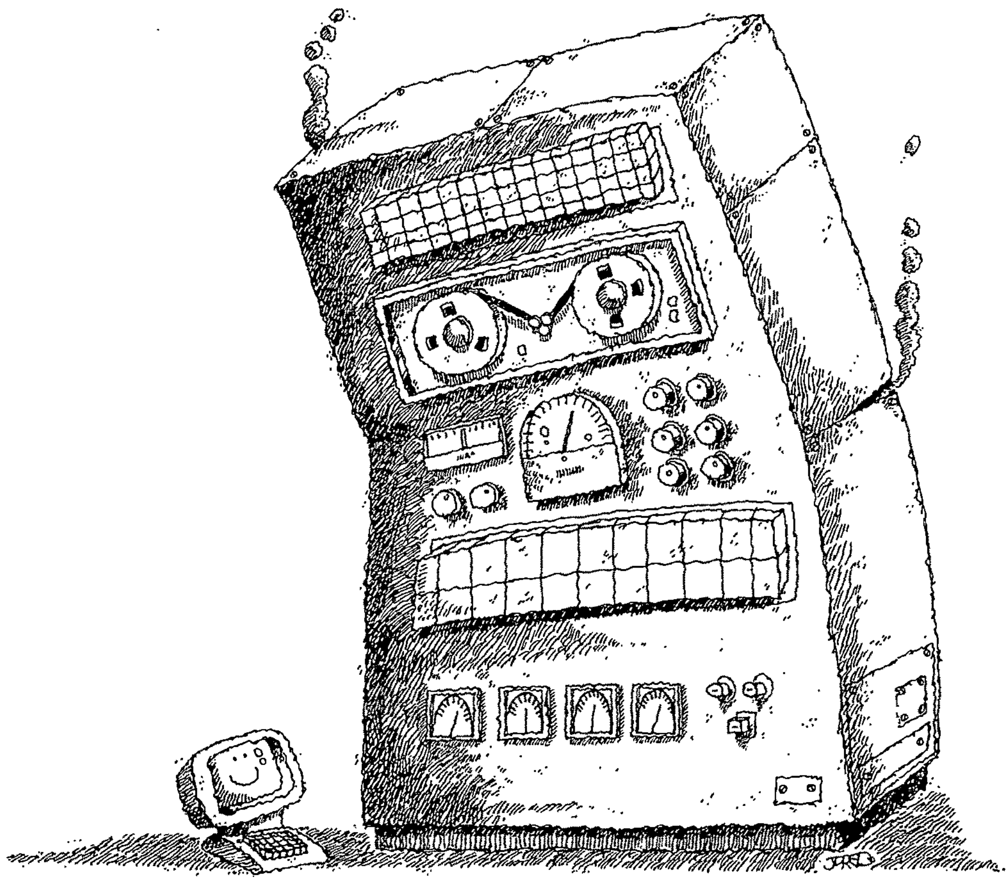
Today's business community is experiencing a growing trend to organize employees into small, entrepreneurially oriented groups. It represents a departure by management from function oriented activity to objective oriented activity in order to survive in increasingly volatile markets. Daniel Roman's unprecedented new book, **MANAGING PROJECTS: A Systems Approach,** provides a much-needed perspective on this trend by focusing on management within the project operational structure. The book explores the complete cycle of project activities from the conception of a project through its termination and evaluation.

Based on the author's 30 years experience in teaching and in industry, **MANAGING PROJECTS** will provide expert advice for the newly appointed project manager, as well as invaluable insight for the experienced manager by covering the range of project activities in a variety of organizational settings. In addition, the book's comprehensive treatment of project management makes it an invaluable textbook for an area rapidly growing in importance.

October 1985,
442 pages,
0-444-00966-3,
Cloth, $34.95
(Dfl. 135 outside North America)



Send orders to:
In North America:
Elsevier Science Publishing Co., Inc.
P.O. Box 1663, Grand Central Station, New York, NY 10163
In the rest of the world:
Elsevier Science Publishers, Direct Mail Department
P.O. Box 211, 1000AE Amsterdam, The Netherlands
Note: Prices subject to change without notice. In NY State please add applicable sales tax.
10/85 NH 167



Bigger does not always mean better.

The first computers were so big they filled entire rooms, and had to be shut down periodically because of overheating.

Things got better when things got smaller. And that's something to consider if you're looking for a publisher.

Now, we at South-Western Publishing would never suggest you select a publisher solely based on size. We would, however, suggest you select a publisher based on quickness of reaction to your needs, personalized service, and flexibility. Because your book isn't like everyone else's. So it shouldn't be treated that way.

So call us at (513) 527-6384. And remember that while some of the big, prodigious publishers may look impressive, they might not respond to your needs the way we will. The important thing about a publishing company is not how big it is, but how big it thinks.

SOUTH-WESTERN
COLLEGE DIVISION

Better Business Begins in the Classroom...

with texts from Merrill

New for 1986 in Management...

UNDERSTANDING ORGANIZATIONAL BEHAVIOR: A Managerial Viewpoint
Robert E. Callahan, Seattle University
C. Patrick Fleenor, Seattle University
Harry R. Knudson, University of Washington
***#20198-5:** Casebound, 508 pages; Instructor's Manual, Test Bank*

UNDERSTANDING ORGANIZATIONAL BEHAVIOR: A Casebook
Harry R. Knudson, C. Patrick Fleenor, Robert E. Callahan
***#20481-X:** Paperback, 280 pages*

PERSONNEL: Management of Human Resources, Second Edition
Michael R. Carrell and Frank E. Kuzmits,
both, University of Louisville
***#20466-6:** Casebound, 712 pages; Instructor's Manual with Test Bank*

EXPERIENTIAL EXERCISES IN PERSONNEL/HUMAN RESOURCE MANAGEMENT, Second Edition
Frank E. Kuzmits
***#20424-0:** Paperback, 256 pages; Instructor's Manual*

FUNDAMENTALS OF STRATEGIC MANAGEMENT
Fred R. David, Mississippi State University
***#20551-4:** Casebound, 868 pages; Instructor's Manual, Test Bank*

STRATEGIC MANAGEMENT SIMULATION
Kenneth B. Olm, George G. Eddy, Eugene Krus, Jr.,
all, University of Texas at Austin
***#20538-7:** Paperback, 90 pages; Disk*

New for 1986 in Marketing...

RETAILING, Second Edition
Dale M. Lewison, University of Akron
M. Wayne DeLozier, University of South Carolina
***#20474-4:** Casebound, 768 pages; Instructor's Manual, Test Bank, Student Guide, Transparency Masters*

RETAILING: Cases and Applications, Second Edition
Dale M. Lewison and M. Wayne DeLozier
***#20558-1:** Paperback, 276 pages; Instructor's Manual*

DIRECT MARKETING
Herbert Katzenstein and William S. Sachs,
both, St. John's University
***#20305-8:** Casebound, 480 pages; Instructor's Manual with Transparencies, Test Bank*

New for 1986 in Finance...

PERSONAL FINANCE: An Integrated Planning Approach
Bernard J. Winger and Ralph R. Frasca, both, University of Dayton
***#20479-8:** Casebound, 672 pages; Instructor's Manual and Study Guide (by Paul R. Allen, Sam Houston State University), Test Bank, Transparency Masters, Software*

New for 1986 in Information Systems...

BUSINESS APPLICATIONS SOFTWARE FOR THE IBM PC®
Lon Ingalsbe, Portland State University
***#20476-3:** Paperback, 350 pages; Instructor's Manual*

Also available by the same author:
LOTUS 1-2-3® FOR THE IBM PC,
dBASE II® FOR THE IBM PC,
WORD STAR® FOR THE IBM PC
(each with Instructor's Manual)

COMPUTERS: An Introduction
Donald D. Spencer
***#20559-X:** Paperback, 352 pages; Instructor's Manual, Computerized Test Bank, Transparencies, Software, Study Guide*

SYSTEMS ANALYSIS AND DESIGN: A Case Study Approach, Second Edition
Robert J. Thierauf, Xavier University
***#20229-9:** Casebound, 672 pages; Instructor's Manual*

INTRODUCING PASCAL: Workbook and Exercises
Barbara Sparks Harris and Christine Benedyk Kay,
both, DeVry/Chicago
***#20454-2:** Paperback, 194 pages; Instructor's Manual*

THE ILLUSTRATED COMPUTER DICTIONARY, Third Edition
Donald D. Spencer
***#20528-X:** Paperback, 320 pages. **#20529-8:** Casebound, 320 pages*

Also new for 1986...

THE LEGAL ENVIRONMENT OF BUSINESS: Commerce and Public Policy
Charles R. McGuire, Illinois State University
***#20325-2:** Casebound, 800 pages; Instructor's Manual with Test Bank, Student Guide*

FUNDAMENTALS OF BUSINESS STATISTICS
John E. Hanke and Arthur G. Reitsch,
both, Eastern Washington University
***#20333-3:** Paperback, 400 pages; Instructor's Manual with Test Bank*

We've many more excellent texts for your courses...
WRITE FOR YOUR COPY OF OUR NEW 1986 CATALOG!

To receive your copy of our catalog, or to request an examination copy of any of these texts for adoption consideration, write: Merrill Publishing Company, ATTN: College Advertising, 1300 Alum Creek Drive, Box 508, Columbus, Ohio 43216. Or, call Chris Cole, toll-free, 1-800-848-1567; in Ohio, call 1-614-890-1111.

MERRILL
PUBLISHING COMPANY
A Bell & Howell Company

Two New Editions from Rue & Byars

Just Published

MANAGEMENT
Theory and Application
Fourth Edition
All New Content!
All New Package!
(Includes Color Transparencies)
See what MANAGEMENT IN ACTION is all about!

Leslie W. Rue, Georgia State University, and
Lloyd L. Byars, Atlanta University

Published Fall 1985

SUPERVISION
Key Link to Productivity
Second Edition
We've added Experiential Exercises to
the New Leader in Supervision
paperbound

Examination copies for adoption consideration
available on request. Please write:

IRWIN
1818 Ridge Road Homewood, IL 60430

Irwin Management Texts for the Way You Teach.

ESSENTIAL REFERENCES
From JOHN WILEY & SONS

BANK STRATEGIC MANAGEMENT AND MARKETING
DEREK F. CHANNON, Manchester Business School
1-90383-3 248 pp. 1986 $31.95

CASES IN BANK STRATEGIC MANAGEMENT AND MARKETING
DEREK F. CHANNON, Manchester Business School
1-90884-3 480 pp. 1986 $49.95

MAKING DECISIONS
Second Edition
D.V. LINDLEY
1-90803-7 207 pp. 1985 $27.95

QUANTITATIVE METHODS IN MANAGEMENT
Case Studies of Failures and Successes
Edited by C.B. TILANUS, Eindhoven University of Technology, O.B. DE GANS, Delft University of Technology, and J.K. LENSTRA, Centre for Mathematics and Computer Science, Amsterdam
1-90841-X 304 pp. 1986 $31.95

OPERATIONS MANAGEMENT IN SERVICE INDUSTRIES AND THE PUBLIC SECTOR
Text and Cases
CHRIS VOSS, University of Warwick, COLIN ARMISTEAD, Manchester Polytechnic, BOB JOHNSTON, University of Warwick, and BARBARA MORRIS, Sheffield City Polytechnic
1-90801-0 318 pp. 1985 $24.95

OFFICE AUTOMATION
A Social and Organizational Perspective
R.A. HIRSCHHEIM, Oxford University
1-90909-2 336 pp. 1986 $29.95

OPERATIONAL RESEARCH
D.J. WHITE, University of Manchester
1-90717-0 361 pp. 1985 $42.95

BRITAIN AND THE MULTINATIONALS
JOHN M. STOPFORD, London Business School, and LOUIS TURNER, Royal Institute of International Affairs
1-90839-8 304 pp. 1986 $24.95

QUANTITATIVE METHODS FOR PUBLIC ADMINISTRATION IN DEVELOPING COUNTRIES
SCOTT C. IVERSON, University of Washington
1-90755-3 167 pp. 1985 $34.95

SURVEY DESIGN AND ANALYSIS
F.R. JOLLIFFE, Brunel University
0-20272-6 178 pp. 1986 $64.95

To order, write to Julie Harrington, Dept. 6-0574:
JOHN WILEY & SONS, INC.
605 THIRD AVENUE, NEW YORK, N.Y. 10158
OR CALL TOLL FREE: 1-800-526-5368.
In New Jersey, call collect: (201) 342-6707. In Canada: 22 Worcester Road, Rexdale, Ontario M9W 1L1.
Order Code #6-0574.
Prices subject to change and higher in Canada.

6-0574

ACADEMY OF MANAGEMENT
1986 REGIONAL MEETINGS

For more information or submissions contact the regional chairperson listed.

SOUTHWEST—March 12-15, Hyatt Regency, Dallas, Texas.
Linda Calvert, (713) 488-9420
Department of Management
University of Houston-Clear Lake
2700 Bay Area Blvd.
Houston, Texas 77058

WESTERN—March 20-22, MGM Grand, Reno, Nevada.
Patrick Connor, (503) 370-6444
Graduate School of Management
Willamette University
900 State St.
Salem, Oregon 97301

MIDWEST—April 17-19, Park Terrace Airport Hilton, St. Louis, Missouri.
Kenneth W. Wexley, (517) 353-5415
Department of Management
228 Eppley
Michigan State University
East Lansing, Michigan 48824

EASTERN—May 7-10, The Hershey Hotel, Philadelphia, Pennsylvania.
Michael Kavanagh, (518) 442-4929
School of Business
State University of New York at Albany
Albany, New York 12222

SOUTHERN—November 12-15, Atlanta Hilton, Atlanta, Georgia.
W. Alan Randolph, (803) 777-5966
College of Business Administration
University of South Carolina
Columbia, South Carolina 29208

ANNOUNCEMENTS

National Academy

Academy of Management 46th Annual Meeting
August 13-16, 1986

Hyatt Regency
Chicago, Illinois

For further information contact:

Don Hellriegel
Vice President-Program Chair
Department of Management
Texas A&M University
College Station, Texas 77843
(409) 845-1724

Membership

Membership applications may be obtained by contacting:

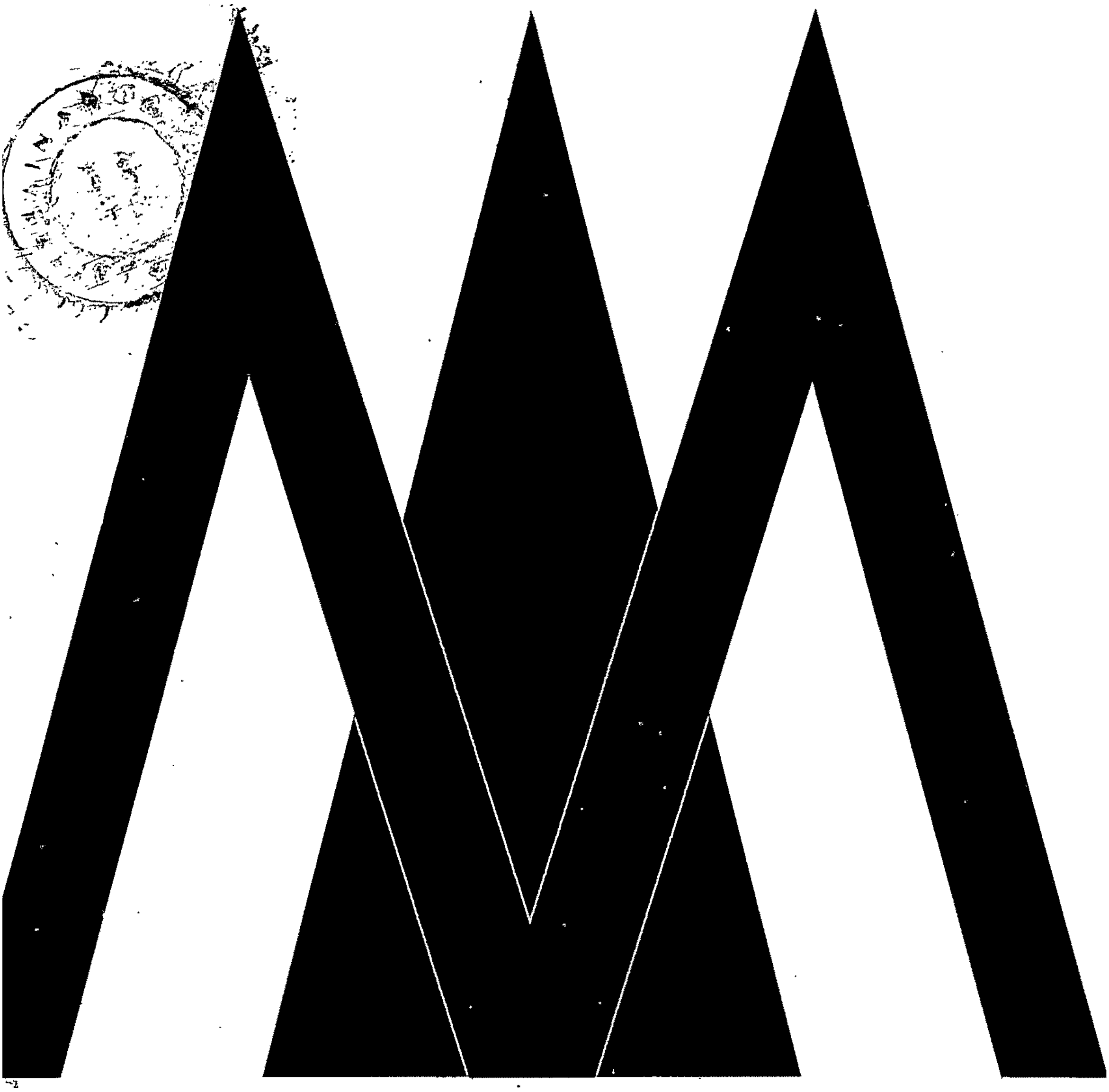
Joseph P. Yaney
Director of Membership
College of Business Administration
Texas Tech University
Lubbock, Texas 79409-4320

Proceedings

Copies of the *Proceedings* of the National Conventions may be ordered from:

Walter B. Newsom
Business Manager
Academy of Management Publications
Mississippi State University
P.O. Box KZ
Mississippi State, Mississippi 39762

Academy of Management JOURNAL



LUME 29 NUMBER 2 JUNE 1986

ACADEMY OF MANAGEMENT

BOARD OF GOVERNORS

PRESIDENT	Fred Luthans, *University of Nebraska*
PRESIDENT-ELECT & PROFESSIONAL DIVISIONS COORDINATOR	Richard M. Steers, *University of Oregon*
VICE PRESIDENT & PROGRAM CHAIRPERSON	Don Hellriegel, *Texas A&M University*
SECRETARY-TREASURER & BUSINESS MANAGER, ACADEMY OF MANAGEMENT PUBLICATIONS	Walter B. Newsom, *Mississippi State University*
PAST PRESIDENT	Kathryn M. Bartol, *University of Maryland*
EDITOR, ACADEMY OF MANAGEMENT EXECUTIVE	W. Warner Burke, *Columbia University*
EDITOR, ACADEMY OF MANAGEMENT JOURNAL	Janice M. Beyer, *State University of New York-Buffalo*
EDITOR, ACADEMY OF MANAGEMENT REVIEW	Orlando Behling, *Bowling Green State University*
DIRECTOR OF MEMBERSHIP	Joseph P. Yaney, *Texas Tech University*
REPRESENTATIVES-AT-LARGE	Joseph A. Alutto, *State University of New York-Buffalo* Michael Beer, *Harvard Business School* Richard B. Chase, *University of Southern California* John M. Ivancevich, *University of Houston* Edwin L. Miller, *University of Michigan* Jeffrey Pfeffer, *Stanford University* Karlene H. Roberts, *University of California-Berkeley* Mary Ann Von Glinow, *University of Southern California*

CHAIRPERSONS OF PROFESSIONAL DIVISIONS

BUSINESS POLICY & PLANNING	David B. Jemison, *Stanford University*
CAREERS	Donald D. Bowen, *University of Tulsa*
HEALTH CARE ADMINISTRATION	Michael Abelson, *Texas A&M University*
INTERNATIONAL MANAGEMENT	Bhal Bhatt, *University of Toledo*
MANAGEMENT EDUCATION & DEVELOPMENT	William J. Heisler, *Newport News Shipbuilding*
MANAGEMENT HISTORY	Mariann Jelinek, *Case Western Reserve University*
MANAGERIAL CONSULTATION	Linda M. Calvert, *University of Houston-Clear Lake*
ORGANIZATION & MANAGEMENT THEORY	Gerald R. Salancik, *University of Illinois*
ORGANIZATIONAL BEHAVIOR	Randall B. Dunham, *University of Wisconsin*
ORGANIZATIONAL COMMUNICATION	Susan A. Hellweg, *San Diego State University*
ORGANIZATIONAL DEVELOPMENT	Marshall Sashkin, *National Institute of Education*
PERSONNEL/HUMAN RESOURCES	Marc Wallace, *University of Kentucky*
PRODUCTION/OPERATIONS MANAGEMENT	Kathleen A. Johnson, *University of Louisville*
PUBLIC SECTOR	Barry Bozeman, *Syracuse University*
SOCIAL ISSUES IN MANAGEMENT	Frederick D. Sturdivant, *Ohio State University*
WOMEN IN MANAGEMENT	V. Jean Ramsey, *Western Michigan University*

OTHER OPERATIONS

ACADEMY HISTORIAN & ARCHIVIST	Charles D. Wrege, *Rutgers University*
NEWSLETTER EDITOR	Rose Knotts, *North Texas State University*
PLACEMENT CHAIRPERSON	James B. Townsend, *Kansas State University*
SIGMA IOTA EPSILON	B.J. Hodge, *Florida State University*
PROCEEDINGS EDITORS	John A. Pearce II and Richard B. Robinson, Jr., *University of South Carolina*

Academy of Management JOURNAL

EDITOR	Janice M. Beyer, *State University of New York at Buffalo*
CONSULTING EDITORS	Richard Mowday, *University of Oregon* Manuel London, *AT&T Communications* Carl Zeithaml, *Texas A&M University*
EDITORIAL BOARD	Harold L. Angle, *University of Minnesota* Achilles A. Armenakis, *Auburn University* Richard Arvey, *University of Minnesota* Susan Ashford, *Dartmouth College* Stephen R. Barley, *Cornell University* Jean M. Bartunek, *Boston College* Barry Baysinger, *Texas A&M University* Max H. Bazerman, *Massachusetts Institute of Technology* Richard S. Blackburn, *University of North Carolina* Joel Brockner, *Columbia University* Terry Connolly, *University of Arizona* Robert D. Dewar, *Northwestern University* Jane E. Dutton, *New York University* Cecilia M. Falbe, *Baruch College, CUNY* Daniel C. Feldman, *University of Florida* Peter Feuille, *University of Illinois* Charles Fombrun, *University of Pennsylvania* Jeffrey D. Ford, *Ohio State University* Louis W. Fry, *University of Washington* Cynthia V. Fukami, *University of Denver* William Glick, *University of Texas* Robert Golembiewski, *University of Georgia* Donald C. Hambrick, *Columbia University* Kathryn R. Harrigan, *Columbia University* Frederic M. Jablin, *University of Texas* William Joyce, *Dartmouth College* Richard J. Klimoski, *Ohio State University* Thomas Kochan, *Massachusetts Institute of Technology* Linda A. Krefting, *Texas Tech University* Frank Krzystofiak, *State University of New York at Buffalo* Gary P. Latham, *University of Washington* Alan D. Meyer, *University of Oregon* Peter Monge, *University of Southern California* Jerry M. Newman, *State University of New York at Buffalo* John Newstrom, *University of Minnesota-Duluth* Jon L. Pierce, *University of Minnesota-Duluth* Jerry Porras, *Stanford University* John E. Sheridan, *Texas Christian University* Harbir Singh, *University of Pennsylvania* James R. Terborg, *University of Oregon* Dave O. Ulrich, *University of Michigan* John F. Veiga, *University of Connecticut* Alan R. Wilkins, *Brigham Young University* Gary A. Yukl, *State University of New York at Albany* Ray Zammuto, *University of Colorado-Denver* Dale Zand, *New York University*
INDEX EDITOR	Nancy G. Dodd, *University of Hawaii-Hilo*
BUSINESS MANAGER	Walter B. Newsom, *Mississippi State University*
ADVERTISING DIRECTOR	Donna Ledgerwood, *North Texas State University*
PRODUCTION EDITOR	Steven Gomes, *Buffalo, New York*
COPY EDITOR	Seph Weene, *Ithaca, New York*
EDITORIAL ASSISTANT	Denise Fredericks

INFORMATION FOR CONTRIBUTORS

The *Academy of Management Journal* publishes articles in fields of interest to members of the Academy of Management. These fields of interest are reflected in the divisions and interest groups of the Academy listed on the inside front cover.

The *Journal* publishes only original, empirical research as articles or research notes. Manuscripts are considered with the understanding that their contents have not been published and are not under consideration elsewhere. In its articles, the *Journal* seeks to publish reports of research that develops, tests, or advances management theory and practice. All types of empirical methods—quantitative, qualitative, or combinations—are acceptable. The *Journal* does not publish purely conceptual or review articles; these are published in the *Academy of Management Review*. Atheoretical exploratory or survey research, methodological studies, replications or extensions of past research, and commentaries with new empirical content are also of interest for publication as research notes if they make an important contribution to knowledge relevant to management.

Articles and research notes should be written so they are understandable and interesting to all members of the Academy. The contributions of research in all specialized areas to *general* management theory and practice should be made evident. Specialized argot and jargon should be translated into terminology in general use within the fields of management.

Articles should also be written as concisely as possible without sacrificing meaningfulness or clarity of presentation. To save space, tables should be combined and data should be presented in the text wherever possible. Manuscripts submitted for publication as articles should not ordinarily exceed 30 double-spaced typewritten pages, including tables. Manuscripts submitted for publication as research notes should not exceed ten double-spaced typewritten pages, including tables. Everything, including tables, in submitted manuscripts should be typed in double-spaced format on one side of the page. Manuscripts prepared on computers should be printed on letter-quality printers or, if other printers are used, in double-strike or enhanced print.

Decisions regarding publication of submitted manuscripts are based upon the recommendations of members of the Editorial Review Board and other qualified reviewers. All articles and research notes published in the *Journal* are subjected to the same anonymous review process. Reviewers evaluate manuscripts on their significance to the field, their conceptual adequacy, their technical adequacy, their appropriateness, and their clarity of presentation. Reviewers' comments are made available to authors.

Submissions should be sent to Dr. Janice M. Beyer, Editor, *Academy of Management Journal*, School of Management, Jacobs Management Center, State University of New York, Buffalo, New York 14260.[1] Manuscripts should be prepared in accordance with the *AMJ Style Guide for Authors*, which is published periodically and is also available from the editor. Contributors should submit four copies of papers, retaining the original for their files. The *Journal* does not return manuscripts unless requested to do so.

---

[1] See *From the Editor* on p. 201 for new address to be used after August 1, 1986.

Academy of Management JOURNAL

CONTENTS

Volume 29, Number 2, June 1986

From the Editor 201

Blau's Theory of Structural Differentiation Revisited: A Theory of Structural Change or Scale?
John B. Cullen, Kenneth S. Anderson, and Douglas D. Baker 203

Tactics of Implementation
Paul C. Nutt 230

Socialization Tactics, Self-Efficacy, and Newcomers' Adjustments to Organizations
Gareth R. Jones 262

Outcomes of Autonomous Workgroups: A Long-Term Field Experiment
Toby D. Wall, Nigel J. Kemp, Paul R. Jackson, and Chris W. Clegg . . 280

Effects of Team Building and Goal Setting on Productivity: A Field Experiment
Paul F. Buller and Cecil H. Bell, Jr. 305

Environments as Moderators of the Relationship between Strategy and Performance
John E. Prescott 329

Multi-Objective Assessment of Effectiveness of Strategic Planning: A Discriminant Analysis Approach
Vasudevan Ramanujam, N. Venkatraman, and John C. Camillus . . 347

Layoffs, Equity Theory, and Work Performance: Further Evidence of the Impact of Survivor Guilt
Joel Brockner, Jeff Greenberg, Audrey Brockner, Jenny Bortz, Jeanette Davy, and Carolyn Carter 373

Clerical Employees' Perceptions of Interorganizational Career Opportunities
Jacqueline Landau and Tove Helland Hammer 385

RESEARCH NOTES

Effects of Extrinsic Reward on Intrinsic Motivation: A Field Experiment
Paul C. Jordan 405

Influence of Task Interdependence and Number of Poor Performers on Diagnoses of Causes of Poor Performance
Karen A. Brown and Terence R. Mitchell 412

Additive Information Processing and the Relationship between Expectancy of Success and Motivational Force
Adrian Harrell and Michael Stahl 424

Published by the Academy of Management four times a year in March, June, September, and December. Subscription rates for U.S., Mexico, and Canada: $38 for one year, $70 for two years, $96 for three years. Rates for overseas: $42 for one year, $78 for two years, $108 for three years. All payable in U.S. dollars.

Subscriptions and orders for back issues should be sent to the Business Manager, Dr. Walter B. Newsom, P.O. Drawer KZ, Mississippi State University, Mississippi State, Mississippi 39762.

Manuscripts and other material of an editorial nature should be sent to Dr. Janice M. Beyer, Editor, *Academy of Management Journal*, New York University, 100 Trinity Place, New York, New York 10006.

Copies of articles published in the *Journal* may be made for instructional, non-commercial use. The Academy of Management does not require prior clearance and makes no charge for this permission.

Inquiries related to advertising should be addressed to Donna Ledgerwood, College of Business Administration, North Texas State University, Denton, Texas 76203.

Typesetting by Repro Services, Inc., Decatur, Georgia; presswork, binding, and mailing by Darby Printing Company, Atlanta, Georgia.

Academy of Management Journal, Publication No. 900920.

Second class postage paid at Mississippi State, MS, and additional mailing offices.

POSTMASTER—Send Form 3579 to Academy of Management *Journal*, P.O. Drawer KZ, Mississippi State University, Mississippi State, MS 29762.

ISSN, 0001-4273

FROM THE EDITOR

Because I will be joining the faculty at New York University in the fall, the *Journal* office will be moved in midsummer. After August 1, please send all manuscripts, revisions, and any other correspondence to:

Janice M. Beyer, Editor
Academy of Management Journal
Graduate School of Business Administration
New York University
100 Trinity Place
New York, NY 10006

I'd like to take this opportunity to thank Dean Joseph Alutto and my colleagues in the School of Management at the State University of New York at Buffalo for the splendid support they have provided to me and to the *Journal* since my editorship began. Not only have they been willing to absorb some of my teaching load and provide resources for the *AMJ* office beyond what the Academy support provides, but my colleagues have been most generous and patient in providing me with advice and expertise when I asked for it—often at short notice and my convenience, not theirs. I'd therefore like publicly to acknowledge their valuable help and support and let them know how much I have appreciated it.

J.M.B.

FORTHCOMING IN THE SEPTEMBER ISSUE

ARTICLES

Glick, Jenkins, & Gupta	Task Characteristics and Attitudes
Ashford	Feedback Seeking and Adaptation
Spencer	Employee Voice and Retention
Wexley & Baldwin	Posttraining Strategies and Positive Transfer
Binning, Zaba, & Whattan	Biasing Effects of Performance Cues
Joyce	A Social Experiment on Matrix
Singh	Performance, Slack, and Risk-Taking in Decision Making

RESEARCH NOTES

Tsui & Barry	Interpersonal Affect and Rating Errors
Oldham, Kulik, Stepina, & Ambrose	Situational Factors and Comparative Referents
Jung, Dalessio, & Johnson	Stability of Factor Structure of JDI
Brockner & Hess	Self-Esteem and Task Performance in Quality Circles
McKinley, Cheng, & Schick	Perceptions of Resource Criticality
Kinicki, Heyl, & Callerman	Validity of Cox-Zmud-Clark MRP Instrument

© Academy of Management Journal
1986, Vol. 29, No. 2, 203–229.

BLAU'S THEORY OF STRUCTURAL DIFFERENTIATION REVISITED: A THEORY OF STRUCTURAL CHANGE OR SCALE?

JOHN B. CULLEN
KENNETH S. ANDERSON
University of Nebraska
DOUGLAS D. BAKER
Washington State University

The major objective of this study was to determine if Blau's (1970) theory of structural differentiation better captures differences in scale between organizations or structural changes within organizations. Data gathered on 134 Ph.D.-granting universities over nine years permitted us to compare results of within-organizations tests of Blau's theory with those from both cross-sectional and longitudinal between-organization tests. The findings suggest that the theory's power lies in explaining structural differences between organizations rather than in explaining structural change.

In an area that has been called a "theoretical wasteland" (Kimberly, 1976: 573), Blau's (1970) theory of structural differentiation in organizations provided a major theoretical statement concerning the effects of organizational size on structure. This theory has inspired considerable research and debate (e. g., Mileti, Gillespie, & Haas, 1977; Miller & Conaty, 1980). The theory has also influenced many studies of formal organizational structure not directly concerned with testing the theory's propositions, probably because of the emphasis given to organizational size and the causal priority the theory grants to size over other dimensions of structure. The large number of citations referencing Blau's initial theoretical paper evidences the extent of his theory's influence. *Current Contents/Social and Behavioral Sciences* recently designated the 1970 article a "citation classic" (Blau, 1983). Moreover, interest has not waned; citations of the initial articles on the theory (Blau, 1970; Blau & Schoenherr, 1971) continue to be numerous more than a decade after their publication (Institute for Scientific Information, 1984).

In spite of the effect of the theory of structural differentiation on many current studies of organizational structure, even recent empirical tests of the theory have used cross-sectional data (Beyer & Trice, 1979). This situation is surprising since a small but growing body of general organizational research

---

We would like to thank three anonymous reviewers for their comments on an earlier draft of this paper. All errors remain the responsibility of the authors.

has shown that cross-sectional differences in structure found among organizations do not necessarily reflect processes of structural change (Freeman & Hannan, 1975). Although a few tests of Blau's theory have used longitudinal data (Meyer, 1972), they merely used change-over-time as a variable in cross-sectional comparisons among organizations (Ford, 1980a). There are virtually no within-organization, time series investigations of the basic relationships that Blau hypothesized. Thus, although his propositions concerning the relationships among size of organization, differentiation, and size of administration continue to influence organizational scholars, especially those concerned with organizational size, the proper sphere of the theory of structural differentiation remains undefined. To what extent is it a theory of *scale* among organizations? To what extent is it a theory of *change* within organizations?

In light of longitudinal research on organizational structure, it seems increasingly clear that both theoretical and empirical work should specify structural relationships that occur because of: (1) differences among organizations in the scales of their operations, (2) differences among organizations in the degrees of change in their structures, and (3) changes within organizational structures over time. Because Blau's work is perhaps the most influential theoretical consideration of organizational size, differentiation, and size of administrative components, it seems appropriate to examine it in terms of these three issues.

Thus, this study had two objectives. The first objective was to test the fundamental aspects of the theory using longitudinal data and a within-organization design. We used within-organization analyses to examine the theory's proposed relationships among the structural dimensions of size, differentiation, and administrative components over time and for individual organizations. The second objective was to compare these results with two types of models: (1) models like those typically reported for studies using cross-sectional data, and (2) models like those reported for studies using longitudinal data and between-organization designs.

BLAU'S THEORY OF STRUCTURAL DIFFERENTIATION

In his formal theory of differentiation, Blau (1970) inferred two generalizations from cross-sectional data on 53 governmental agencies and their local branches. The generalizations were: "Increasing size generates structural differentiation along various dimensions at decelerating rates" (1970: 204) and "Structural differentiation in organizations enlarges the administrative component" (1970:213). Blau operationally defined organizational size as number of employees, differentiation as number of formal structural components such as levels or divisions, and an administrative component as the number of staff personnel who provide support services for an organization.

Blau derived numerous propositions from the two generalizations. These propositions hypothesized at least three basic relationships: (1) increasing organizational size reduces the relative size of an administrative component,

(2) increasing organizational size increases differentiation, and (3) increasing differentiation increases the relative size of an administrative staff. Blau (1972) later expressed the proposed empirical relationships in path analytic form, as shown in Figure 1.

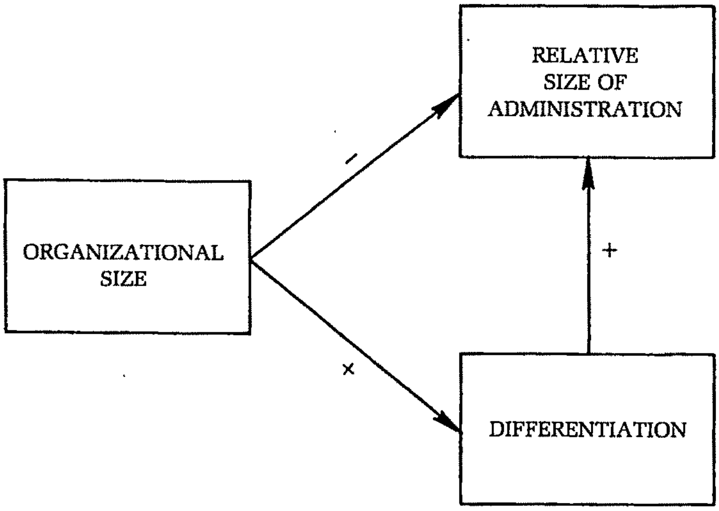
Although Blau used cross-sectional data to derive his theory, the focus of his basic generalizations on *changes* in structural components implied a temporal dimension to organizational variability. However, the vast majority of research related to this theory has used cross-sectional data and examined only differences among organizations. As the literature reviewed in this study shows, the theory of structural differentiation is useful for explaining differences in scale among organizations. Yet we know little concerning its applicability to processes of change within organizations. As a result, the basic relationships of the theory still need to be examined over time for individual organizations if Blau's theory is to be considered a viable explanation of processes of organizational change.

Issues Emerging from the Study of Change

Research that examines Blau's basic hypotheses over time, especially with a within-organization design, should consider two issues that neither he nor the few longitudinal tests of his theory have addressed.

The first issue is the lag structure, or the time period it takes relevant variables to respond to each other. Cullen and Baker (1984) demonstrated that the responsiveness of changes in size of administration to changes in enrollment varied widely in a random sample of colleges and universities. These findings were similar to those in other fields where within-subject time series analyses are more common (e. g., Assmus, Farley, & Lehmann,

FIGURE 1
Simplified Model of the Theory of Structural Differentiation



1984). Little is known about how long it takes differentiation to respond to changes in size or how long it takes administration to respond to changes in differentiation. More than just an empirical issue, investigation of lag structures addresses an unspecified theoretical point in Blau's work. All theories of change need to specify when one variable will respond to changes in another variable. Lacking such specification, any theory becomes irrefutable since hypothesized relationships will nearly always occur at some unspecified time in the future.

The second issue is how the theory of structural differentiation should be interpreted for declining as well as growing organizations. Although its originator phrased the theory's generalizations and propositions in terms of growth, the derivation of the theory of structural differentiation from cross-sectional data relies on an assumption concerning the results of earlier changes in structure. It implies that changes—growth, decline, or both—result in similar combinations of structural characteristics at particular levels of size at any time. Stated otherwise, "[organizational] levels represent the sum of all change up to the point of measurement" (Dewar & Hage, 1978: 112). As a result, Blau's theory at least implies hypotheses concerning decline as reversal of growth. If growth hypotheses are reversed, the following situations are necessary during decline in order for the sum of differences created by growth or decline to produce cross-sectional relationships consistent with the theory.

Since relationships between size and differentiation and size and administration are not proportional during growth, they should not be proportional during decline. Blau's theory states that during growth, decentralization occurs at a decelerating rate because the average size of the subunits increases (Blau, 1970: 206-207). This hypothesis says that there is absolutely more but relatively less differentiation with an increase in size. On the basis of this more general hypothesis concerning differentiation, Blau also hypothesized that the relative size of an administrative staff declines with increases in an organization's size (Blau, 1970: 210). Because it is based on cross-sectional data, the theory of structural differentiation presumes that a decline returns an organization to an earlier state. Therefore, if proportionately larger administrative staffs are expected for smaller organizations, the theory implies that in decline, administrative size decreases slower than organizational size. Similarly, differentiation should decline proportionately less than size declines. Because Blau did not specify whether the relationship between differentiation and administration is proportional during growth, support for his theory only requires that the form of this relationship be similar in growth and decline.

Although the logic of the theory of structural differentiation has not been tested previously with growing and declining organizations, evidence from both between-organization (Freeman & Hannan, 1975) and within-organization (Ford, 1980a) longitudinal studies challenges this implicit hypothesis that growth and decline produce the same cross-sectional relationships among components of organizational structure.

In response to these issues, the within-organization models of the theory of structural differentiation used in this study (1) considered different lag structures for organization size, and (2) carefully examined differences between growing and declining organizations.

PREVIOUS RESEARCH

The following sections review selected previous research on the major relationships that Blau hypothesized. We differentiated studies reviewed along two dimensions: (1) whether they used longitudinal or cross-sectional data, and (2) whether their research designs were within- or between-organizations. The cross-sectional research we examined focused on two or more of the variables that figure in Blau's theory. Cross-sectional studies necessarily used between-organization designs examining differences among or between organizations. Longitudinal research having within- or between-organization designs examined two or more of the theory's variables over two or more time periods. Between-organization studies used longitudinal data to develop change variables for assessing variance in the theory's variables across samples of organizations; within-organization studies examined variance in these variables over time for single organizations.

Cross-Sectional Studies

Cross-sectional research examining the hypotheses posited by Blau essentially replicated his findings (1970, 1972, 1973). Several investigators have studied the relationship between organization size and structural differentiation. Most have found a positive association. Goldman (1973) found strong relationships between organizational size and three forms of differentiation—spatial, horizontal, and vertical. Mileti and colleagues (1977) observed positive curvilinear relationships between size and vertical and horizontal differentiation and a linear relationship between size and spatial differentiation. Miller and Conaty (1980) pooled data from three different sources: (1) federal and state agencies studied by Mileti and colleagues (1977), (2) department stores studied by Goldman (1973), and (3) an original heterogeneous sample contributed by Miller. They concluded that size is related positively to differentiation. In a study attempting a direct replication of the theory of structural differentiation's hypothesis concerning differentiation, Beyer and Trice (1979) found size to have greater effects on division of labor, vertical differentiation, and horizontal differentiation than did Blau and Schoenherr (1971).

Among the relationships postulated in the theory, the effects of size on administration have received the most attention; this was true even before Blau developed his theory. He hypothesized a positive correlation between organizational size and the absolute size of the administrative component and a negative relationship between organizational size and the relative size of the administrative component (Blau, 1970: 208). In their review of the literature, Jackson and Morgan (1978: 238) found unqualified support for a positive relationship between the absolute size of both components. Bedeian's (1980) review of 26 cross-sectional studies with data on organizational size and administrative ratios reported that 18 studies observed a negative relationship between organizational size and relative size of administration. Gillespie and Mileti (1976) concluded from their literature review that, with size controlled, there are positive associations between the ratio of administrators to organizational size and both vertical and horizontal differentiation. All of these findings are consistent with the theory of structural differentiation.

Between-Organization Longitudinal Studies

Most longitudinal studies investigating organizational structure employ between-organization designs. Such studies (e. g., Akers & Campbell, 1970; Freeman & Hannan, 1975) use changes observed within individual organizations in comparisons across or between cases or organizations. For example, in a between-organization design with longitudinal data, once an organization's percentage of growth or decline has been measured, the analysis uses this variable—just like any other variable measured at one point in time—for estimating relationships across or between organizations.

A small number of data points on which change can be measured and the variability of patterns of change within organizations make it difficult to determine how well relationships for change variables derived from between-organization research pertain to the same changes within individual organizations (Ford, 1980a: 617). For example, since between-organization studies usually apply one lag time to all variables and to all organizations studied, differences among the lag structures for individual organizations may obscure subtleties of relationships. Similarly, a single estimate of a relationship between change variables for an entire sample is sensitive to outliers in ways that may mask findings. It is possible, for example, that most organizations in a sample could have a relationship over time when a between-organization slope suggested no such relationship.

However, since between-organization studies with longitudinal data do consider change and can contrast processes of growth and decline, it is unfortunate that only a few such studies exist—especially investigating the theory of structural differentiation or its constituent relationships. Using data collected at two points, Meyer (1972) confirmed his hypothesis that the relationship between size and differentiation was unidirectional; he found that size affects differentiation. In contrast, a later study by Dewar and Hage (1978) concluded that organizational size did not affect vertical or horizontal differentiation.

In a six-year study that supported aspects of Blau's theory, Hendershot and James (1972) found a negative association between organizational size (enrollment) and administrative intensity (the ratio of principals and supervisors to teachers) in rapidly growing organizations. However, they found no consistent associations for organizations that grew slowly. Daft and Becker (1980) hypothesized and found that organizational size had a positive relationship with the absolute size of the administrative component.

Freeman and Hannan (1975) examined organizational size and size of the administration for both growing and declining school districts. The relationship in growing districts between district size and supportive components—an inclusive measure of administration—was positive, proportional, and relatively consistent with Blau's theory. Declining organizations removed members from their teaching or direct components faster than from their supportive components. Similarly, Freeman (1979) found a positive relationship between district size and size of the supportive components in both growing and declining districts. Enrollment had a stronger positive effect on supportive components in growing districts than in declining districts.

If the cross-sectional differences found originally in the derivation of Blau's theory represent the results of changes that predated his analysis, the diseconomies in administration associated with small organizations should also be produced in formerly larger organizations moving into the smaller size because of decline. Thus, findings of decreased economies of scale with decline are logically consistent with the theory of structural differentiation, even though Blau did not deal with decline explicitly. However, critics have suggested that highly rational approaches to organizational change, such as Blau's, ignore the possibility that the political power of administrators becomes manifest during decline and is used to maintain the positions of those in power (Hannan & Freeman, 1978).

Without direct study of the politics of decision making in an organization it is difficult to assess the degree to which rationality or political power prevail. Hannan and Freeman's (1978) findings did indicate indirectly that political power operates to increase retention of administrators in decline more than in growth. Ford (1980b) also argued that the relationships between size and various components of structure in growing organizations will not mirror those relationships found in declining organizations. During periods of decline, organizational decision makers may offer employees inducements to stay in an attempt to stem the processes of decline. Ford referred to this as the participation problem (1980b: 592-593). Commitment of an organization's decision makers to its status quo—their resistance to change—may shape relationships in decline. Finally, like Hannan and Freeman (1978), Ford argued that political considerations during decline may lead to an organization's members wishing to maintain their positions, thus limiting the organization's ability to adapt.

Daft and Becker (1980) hypothesized that organizational complexity is related positively to size of the administrative component. They found a positive but weak association between complexity and changes in size of administrative components. Moreover, the relationship was reciprocal, with an additional weak but positive association between size of administrative components and changes in complexity.

Although longitudinal research with between-organization designs has tended to support the major hypotheses of Blau's theory, these studies provided less support than has cross-sectional research.

Within-Organization Longitudinal Studies

A within-organization longitudinal design typically examines relationships over time for single organizations. Data are thus compared within individual organizations rather than between different organizations. Such studies then use regression models to explain variance in a time series for two or more organizational variables. Regressions for single organizations are usually compared descriptively.

In spite of some early within-organization research (Tsouderos, 1955), longitudinal studies of organizational structure with within-organization designs are still rare (Kimberly, 1976). The difficulty of obtaining numerous data points in a time series probably accounts for this scarcity. We know of no within-organization studies of the theory of structural differentiation that have considered organizational size, differentiation, and size of administration simultaneously. However, there is at least one recent within-organization study of organizational size and size of administration.

Ford (1980a) collected data over ten years (1968-77) on 24 public school districts, hypothesizing little convergence between the results of cross-sectional analyses and the results of longitudinal analyses. He concluded that it was "problematic" (1980a: 627) to infer organizational processes from the results of research using cross-sectional designs. Ford used size elasticities—proportional changes over time—to assess administrative economies of scale in growing and declining organizations. For school districts with growing enrollments, he found diseconomies of scale for instructional administrators while approximately 50 percent of the school districts had economies of scale for noninstructional administrators. Most declining school districts added both types of administrators. Although the inverse of Blau's hypothesis of administrative economies of scale would suggest diseconomies during decline, these diseconomies should be slightly greater than proportional in order for the differences in scale between organizations to remain consistent with his theory, which states that small organizations have proportionately more administrators. However, the absolute increases in administrative size during decline that Ford found are diseconomies exceeding those predicted by Blau; they perhaps result from the differential power of administrators (Hannan & Freeman, 1978), or from structural hysteresis (Ford, 1980b).

Summary

Although the investigations cited in the preceding sections represent only some of the research on the major variables of the theory of structural differentiation, they illustrate the basic trends in the findings for three different methodologies (cross-sectional, between-organization with longitudinal data, and within-organization analyses). Most cross-sectional research has tended to confirm the theory's major hypothesized relationships for size, differentiation, and administration. For growing organizations, within- and between-organization longitudinal research has tended to support the hypothesis of administrative economies of scale. In addition, studies that have looked at organizations declining in size have often found support for the inverse of the hypothesis of economies of scale during growth; diseconomies with decline were found. However, these diseconomies have often been more extreme than those suggested by the theory of structural differentiation, with many organizations actually increasing the absolute sizes of their administrative staffs. The few longitudinal studies of differentiation, either as a predictor of administrative size or as a dependent variable to organization size, have been inconclusive in their assessment of Blau's theory.

It is noteworthy that there has been almost no research examining the three major variables of Blau's theory simultaneously and longitudinally. Moreover, those longitudinal investigations that do exist have between-organization designs. Thus, to better understand the processes that Blau and others implicitly hypothesized to take place over time, further longitudinal research on the theory of structural differentiation seems necessary (cf. Meyer, 1971, 1972). Cross-sectional research has already provided fairly strong support for the theory. Results from within-organization longitudinal research—when contrasted with those from cross-sectional and between-organization longitudinal research—can help determine the degree to which the theory applies to changes within individual organizations or to differences among organizations in the relative scales of their structural characteristics. The study reported here does just this by testing selected basic hypotheses of Blau's theory using cross-sectional models, between-organization longitudinal models, and within-organizational models. The within-organization models show internal change processes in individual colleges and universities over a nine year period.

METHODS

Population and Data Source

Four-year universities that granted Ph. D. degrees were chosen as the population of organizations. We collected data on these organizations over nine academic years, from 1975-76 through 1983-84. For simplicity, we will refer to this period as 1975 through 1983. An original population of 251 Ph. D.-granting universities appeared in the 1973 version of the *American Universities and Colleges* (Furness, 1973). Missing data reduced the number of organizations we studied to 134. As a result, there were 1,206 observations or data points (9 years × 134 organizations).

The source of data was the *Yearbook of Higher Education*, volumes 7-15 (Marquis Academic Media, 1975/76-1983/84). Observations began with the 1975-76 volume because the *Yearbook* began reporting data consistently on faculty size during that year. Before that time, the only measure of size available was student enrollment. Since Blau's theory deals with differentiation in organizational workforces, data on faculty size were necessary to provide an appropriate measure.

Although published data sources are invaluable for providing time series data on organizations (Miller & Friesen, 1981), they provide few variables for introducing statistical control. We selected universities that granted Ph.D. degrees because we needed a population that was relatively homogeneous in organizational purpose but varied in size, differentiation, and administration. Public universities represented 55.2 percent of the data set. Relationships reported in the Results section were not changed substantially when controls were introduced for public auspices.

The academic year was used as the period of measurement because universities normally adjust budgets, students, faculty, and administrators yearly.

Variables

In line with the theory of structural differentiation, variables representing organizational size, size of administration, and structural differentiation were included in the study. As did Blau's (1973) study of colleges and universities, this test of his theory focused on academic workforces and academic administrative components.

The measure of structural differentiation was the number of academic departments. In Blau's test of his theory with academic organizations (1973), he used number of academic departments to indicate academic division of labor. Because faculty comprise the primary workforce differentiated into academic departments by universities, the number of faculty was used as the measure of organization size. Full-time equivalent faculty was defined as the number of full-time faculty plus half of all part-time faculty.

Although the *Yearbook* reports data on all types of administrators above the level of department chairpeople, many of these administrators deal with business, finance, or student personnel functions rather than with faculty in academic and teaching activities. Thus, we used the number of full-time academic administrators above the level of department chairpeople to represent size of administration. Academic administrators were defined as those with direct or indirect authority over teaching or research functions. Independent classifications of administrators into academic or nonacademic categories on the basis of titles showed 91.7 percent agreement between two coders. The senior author resolved discrepancies between coders. Academic administrators included presidents or chancellors; all academic deans; and all division heads for instruction, academic affairs, special research or teaching programs, adult or continuing education, and research. Titles for division heads included vice-president or vice-chancellor, dean, and director.

Several arguments supported choosing academic administrators above the level of department chairpeople as the administrative category. First and most important in terms of the theory of structural differentiation, we assumed that academic administrators were the administrative component most directly linked to changes in academic differentiation and to changes in the size of academic workforces. Second, the most detailed data published in the *Yearbook* concerned academic administrators. Third, it is generally better to investigate a homogeneous category of administrators since relationships between sizes of administration and other variables might differ among various possible categories of administration—an issue raised even before the publication of the functional theory of differentiation (Rushing, 1966). Fourth, although the *Yearbook* reported some data for administration in areas such as business and finance, these data alone were not sufficient to test Blau's theory since no data were reported on either size or departmentalization of nonfaculty workforces. Fifth and finally, we examined academic administration above the level of department chairpeople to avoid definitional dependency between size of administration and differentiation; such dependency is a common danger in relating components of organizational structure (Meyer, 1971). Since departmentalization was the measure of horizontal differentiation, adding the number of chairpeople to the size of an administrative staff would have been virtually equivalent to adding number of departments, an independent variable, to size of administration, the dependent variable. The relationship between these variables would then have been biased.

Although Blau's theory does not give theoretical prominence to one type of differentiation, it does hypothesize that differentiation occurs along two basic dimensions—the horizontal and the vertical. Academic departments exhibit two forms of horizontal differentiation—into subunits and into specialized tasks. Previous researchers have noted both positive and negative simple correlations between horizontal and vertical differentiation (Beyer & Trice, 1979; Miller & Conaty, 1980). Positive relationships have sometimes disappeared (Blau & Schoenherr, 1971) when studies have controlled for size. One interpretation of such findings is that one form of differentiation may substitute for the other, especially within given size levels (Blau, 1973: 58), even though size shapes both forms. As a result, an inclusive test of Blau's theory with longitudinal data would require examining vertical as well as horizontal differentiation. Unfortunately, no information was available in the *Yearbook* for vertical differentiation or number of levels in the organizational hierarchy of the university sample.

This lack seemed less of a problem for testing Blau's theory with academic organizations than it would have been for a test using another type of organization. Our basic reason for this assertion is that, perhaps for historical reasons, there is little variance and a restricted range in the number of hierarchical levels in academic organizations. Blau's (1973) sample of universities and colleges showed a mean of 4.4 levels (s. d. = 0.7). Blau (1973) argued that most organizations had four levels (in terms of his measure): (1) faculty–chair, (2) chair–dean, (3) dean–vice-president, (4) vice-president–president (1973: 29-30). He noted further that "very few [organizations] have more than five or fewer than three levels" (1973:30). Given the nature of these data, it was not surprising that his test of the theory of structural differentiation (Blau, 1973: 54-56) explained considerably more variance in college and university departmentalization ($R^2 = 0.73$) than in number of hierarchical levels ($R^2 = 0.36$). Thus, although vertical differentiation in academic organizations is associated with size, it does not seem to be their major form of differentiation.

Procedures

The analyses focused on the three basic relationships shown in Figure 1: (1) organizational size predicting size of administration, with differentiation controlled, (2) organizational size predicting differentiation, and (3) differentiation predicting size of administration, with organizational size controlled. When these three relationships are considered as a path model, two regression equations are necessary to examine them: the regression of differentiation on organizational size, and the regression of administration size on organizational size and differentiation.

Blau (1970: 210, 214) initially tested the hypothesis of an economy of scale in administration for the simple bivariate relationship between size and administration and for the direct effect of size on administration with differentiation controlled. The present study considered only the direct effect, as this was the procedure Blau used in his later work (1973) on colleges and universities. Moreover, this procedure is consistent with the equation suggested by Figure 1 when the figure is considered a path diagram.

Within-organization analyses. We computed separate ordinary least squares (OLS) regressions over time for each organization. For each organization, two basic equations were estimated to test the hypothesized within-organization (time series, t = years) relationships:

$$DIFF_t = SIZE_t \quad , \tag{1}$$

$$ADMIN_t = DIFF_t + SIZE_t \quad .^1 \tag{2}$$

OLS estimates were used because examination of the Durbin-Watson statistic showed only a small degree of autocorrelation. Moreover, parameter estimates are usually unbiased even with autocorrelated disturbances (Ostrom, 1978).

Following Ford (1980a), we did not use inferential statistics since the organizations studied were not a random sample, nor were the observation years sampled randomly over organizational life spans. Regressions over the observation years merely describe yearly changes in the characteristics of the observed individual organizations during the time of the study. We considered Blau's theory confirmed or negated on the basis of two criteria, slope direction and magnitude. This procedure is not inherently biased for or against Blau's theory but only reveals whether the observed organizations behaved consistently with the theory from 1975 through 1983. We make no attempt to generalize to all academic organizations or to the entire life spans of the organizations examined.

For two reasons, we transformed all variables to a proportional scale. First, the theory of structural differentiation considers proportional or relative changes in relationships among variables, such as the administrative economy of scale during growth in size. Second, earlier within-organization

---

[1] *DIFF* = differentiation; *SIZE* = organization size; *ADMIN* = administration size.

studies (Ford, 1980a) have shown the value of using size elasticities to examine the results of proportional changes in size over time. For both independent and dependent variables, we used a logarithmic scale of base 10 to assess the elasticity or responsiveness of a proportional change in one variable to a proportional change in another variable.

Using proportional transformations on both sides of an equation makes a straightforward interpretation possible. A slope of 1 in a regression equation indicates that a change of one percentage point in the independent variable produces a change of one percentage point in the dependent variable. A slope of less than 1 indicates a proportionally smaller change in the dependent variable, and a slope of greater than 1 indicates a proportionally greater change in the dependent variable. For example, a positive slope of less than 1 shows an administrative economy of scale for a growing organization. Such transformations have the additional advantage of showing proportional changes without using potentially suspect ratio variables (Freeman & Kronenfeld, 1973; MacMillan & Daft, 1979). Although many authors, including Blau (1970, 1973), used logarithmic transformations of size in estimating the nonlinear or declining effect of size on untransformed measures of administration and differentiation, theoretical interpretation of regression slopes is more ambiguous with this method than it is when both sides of an equation are measured proportionally.

Growing and declining organizations were analyzed separately, consistent with previous longitudinal investigations (Ford, 1980a; Freeman & Hannan, 1975). An organization was defined as growing if the slope of the regression of organizational size on the yearly observation dates was positive. A negative slope indicated a declining organization. This is essentially the same as comparing the mean size of an organization in the later years of the study with its mean size in the earlier years.

Time lags for administrative changes following changes in organization size vary (Cullen & Baker, 1984; Freeman, 1979). Moreover, researchers in other fields have observed variable lag structures for organizational actions (Assmus, Farley, & Lehmann, 1984). Given these considerations, we examined three lag structures for organizational size in testing Blau's theory: (1) contemporaneous, with organizational size and the dependent variables, differentiation and size of administration, measured over the same years; (2) the one-year lag, with size observed one year earlier than the dependent variables; and (3) the two-year lag, with size observed two years earlier than the dependent variables. The following six equations were computed for each organization:

$$DIFF_{77\text{-}83} = SIZE_{77\text{-}83} \quad , \tag{3}$$

$$DIFF_{77\text{-}83} = SIZE_{76\text{-}82} \quad , \tag{4}$$

$$DIFF_{77\text{-}83} = SIZE_{75\text{-}81} \quad , \tag{5}$$

$$ADMIN_{77\text{-}83} = DIFF_{77\text{-}83} + SIZE_{77\text{-}83} \quad , \tag{6}$$

$$ADMIN_{77\text{-}83} = DIFF_{77\text{-}83} + SIZE_{76\text{-}82} \quad , \tag{7}$$

$$ADMIN_{77\text{-}83} = DIFF_{77\text{-}83} + SIZE_{75\text{-}81} \quad . \tag{8}$$

Because previous research has identified no best lag structure for the effects of organizational size on size of administration (Cullen & Baker, 1984), the within-organization tests of the theory's basic hypotheses reported in this study were based on each organization's best lag model for relationships of administration and differentiation with size. The following criteria were used to determine the best lag models for the relationships of organizational size with size of administration and differentiation. If the R^2 for a size lag (either contemporaneous, one-year, or two-year) was .1 or larger than the other two, we selected that model as the best lag structure for size. If the highest R^2 for any equation was not at least .1 or larger than the other two, we classified the lag structure for organizational size as ambiguous and used the contemporaneous model for analyses. If no variance in the dependent variable was shown, there was no relationship and hence no appropriate lag structure.

Cross-sectional and between-organization longitudinal analyses. The same relationships tested within organizations were tested using cross-sectional models and between-organization longitudinal models. For illustrative purposes, we used 1977 and 1983 data for two cross-sectional tests of the theory of structural differentiation. Between-organization longitudinal models used the same two years as data points for the development of change variables. We chose 1977 and 1983 to facilitate comparisons with the within-organization analyses where variation in administration from 1977 to 1983 represented the final variable in the tests of Blau's theory. Multiple regression equations representing the relationships posited in Figure 1 were used in both types of analyses for 134 organizational subjects.

Variables were the same as those used for the within-organization analyses. However, three different measurement strategies were used to match the measurement approach used in this study's within-organization analyses and measurement procedures found commonly in the literature. First, for both the cross-sectional and between-organization longitudinal analyses, we transformed all variables logarithmically to compare them directly with the within-organization models reported in this study. Second, we used absolute measures of all variables for consistency with previous research that has examined these variables in absolute terms. Third, we used ratio measures of administration size only in the two cross-sectional regressions to achieve consistency with much of the past cross-sectional research including Blau's. The Results section presents more details on these analytical strategies.

RESULTS

Within-Organization Analyses

Of the 134 organizations in the sample, 96 were growing in size and 38 were declining. Table 1 presents a cross-tabulation of the best lag models for organizational size by the two relevant relationships hypothesized in Blau's theory and by growth and decline. No one lag structure was found appropriate for size under conditions of either growth or decline. Separate chi-square

TABLE 1
Percentages of Best Lag Models for Size by Growth and Decline in Size and by Relationships with Size[a]

		Best Lag Models[b]				
Direction of Change	**Relationships**	**0**	**1**	**2**	**?**	**No Variance**
Growth	*SIZE* → *DIFF*	19.8 (19)	15.6 (15)	19.8 (19)	37.5 (36)	7.3 (7)
	SIZE → *ADMIN*	19.8 (19)	10.4 (10)	12.5 (12)	51.1 (49)	6.2 (6)
Decline	*SIZE* → *DIFF*	13.2 (5)	26.3 (10)	26.3 (10)	31.6 (12)	2.6 (1)
	SIZE → *ADMIN*	36.9 (14)	7.8 (3)	21.1 (8)	31.6 (12)	2.6 (1)

[a] Separate chi-square tests were computed as follows: distribution of models by size relationships in growing organizations($\chi^2 = 4.65$, $df = 4$, $p = 0.33$); distribution of models by size relationships in declining organizations ($\chi^2 = 8.25$, $df = 4$, $p = 0.08$); distribution of models by growth/decline for the size → differentiation relationship ($\chi^2 = 4.12$, $df = 4$, $p = 0.39$); distribution of models by growth/decline for the size → administration relationship ($\chi^2 = 7.67$, $df = 4$, $p = 0.10$).

[b] Absolute numbers of organizations are reported in parentheses.

tests on the cross-tabulations of lag models by different relationships within the growth and decline categories and on the cross-tabulations of the same relationships across growth and decline were not significant.

Table 2 shows the distribution of regression slopes for the three relationships hypothesized in the theory of structural differentiation. For relationships in which size was an independent variable, data pertain to the best model, if one existed; data pertain to the contemporaneous model for relationships with ambiguous lag structures for size. Where there was no variance in the dependent variable, there was by definition no support for Blau's theory.

For growing organizations, the theory of structural differentiation hypothesizes absolute increases in size of administration and differentiation that are proportionally less than increases in organizational size. For declining organizations, the theory hypothesizes—albeit implicitly—absolute decreases in size of administration and differentiation that are proportionately less than decreases in organizational size. Since this study computed regression equations separately for all growing or declining organizations and since the slopes of these equations represent the relationships between proportionately transformed variables over time, positive slopes of less than 1 demonstrate support for the theory for either growing or declining organizations.[2] Positive slopes of less than 1 for regressions of differentiation or administration on growth in organizational size indicate situations where these dependent variables grew more slowly than did organizational size. That is, for such organizations, growth resulted in proportionately less differentiation and fewer administrators. Positive slopes of less than 1 in regression equations for declining organizations indicate situations where administration size or differentiation declined more slowly than did organizational size; decline resulted in proportionately more differentiation and more administrators.

Most organizations had positive slopes in the relationships of size and administration and size and differentiation regardless of whether they grew or declined during the study years. However, the forms of the relationships did not always follow expectations derived from Blau's theory.

The regression of differentiation on increases in size resulted in positive slopes for 56.4 percent of the growing organizations. Positive slopes of less than 1 occurred more frequently (39.6%) than did positive slopes greater than or equal to 1 (16.8%), indicating that increases in differentiation often occurred, as expected, less-than-proportionally to increases in size. For organizations declining in size, positive slopes indicated a reduction in horizontal differentiation for 60.5 percent of the organizations. For these organiza-

---

[2] When regressions are estimated between variables over time, neither the direction nor magnitude of the slopes obtained show relationships with time; rather they represent the covariation between the independent and dependent variables. For example, a positive slope for size of administration regressed on organizational size means only that these two variables grow or decline together, regardless of whether organizational size increases or decreases during the time period considered.

TABLE 2
Percentages of Positive and Negative Within-Organization Regression Slopes by Growth and Decline for Relationships in Model

		Slope Distribution[a]				
		Positive		Negative		No
Direction of Change	Relationships	<1	≥1	<1	≥1	Variance
Growth in size	*SIZE* → *DIFF*	39.6 (38)	16.8 (16)	27.0 (26)	9.3 (9)	7.3 (7)
	SIZE → *ADMIN*	31.3 (30)	17.7 (17)	25.0 (24)	19.7 (19)	6.3 (6)
Decline in size	*SIZE* → *DIFF*	50.0 (19)	10.5 (4)	31.5 (12)	5.4 (2)	2.6 (1)
	SIZE → *ADMIN*	26.4 (10)	13.1 (5)	31.5 (12)	26.4 (10)	2.6 (1)
Growth in differentiation	*DIFF* → *ADMIN*	30.2 (28)	18.3 (17)	36.6 (34)	9.6 (9)	5.3 (5)
Decline in differentiation	*DIFF* → *ADMIN*	24.4 (10)	17.0 (7)	12.3 (5)	39.0 (16)	7.3 (3)

[a] Absolute numbers of organizations are reported in parentheses.

tions, differentiation changed most often with relatively less magnitude than did size—a finding consistent with Blau's theory. However, in direct contrast to that theory's expectations, approximately 40 percent of the organizations in this study either deleted departments during growth in size (shown by negative slopes), added subunits during decline (also shown by negative slopes), or did not alter their levels of differentiation in either a growth or decline.

Only the 31.3 percent of the growing organizations with administrative staffs that increased less-than-proportionately to their growth in size showed the specific type of economy of scale predicted by the theory of structural differentiation. Blau (1970, 1972, 1973) predicted consistently that increasing size increases the absolute size of an administrative component while reducing its relative size. The 44.7 percent of the growing organizations with negative slopes for regressions of size of administration on organizational size showed gains in administrative economies, but these were through absolute *reductions* in the sizes of their administrations. Similarly, only 26.4 percent of the declining organizations had positive slopes of less than 1, indicating the expected less-than-proportional reduction in numbers of administrators during decline. The majority of declining organizations (57.9%) had proportionately larger administrative staffs, not because size of administration decreased more slowly than did the decline in size, but because the number of administrators increased absolutely. The negative slopes in Table 2 show this finding.

To better understand the directional effects produced by regressions of size of administration on differentiation, we tested Blau's hypothesis of a positive effect of differentiation on size of administration separately for organizations undergoing growth or decline in differentiation. As with the growth and decline classifications for size, an organization was classified as growing in differentiation if it had a positive regression slope for differentiation on its observation years. A negative slope identified decline. Observation years for differentiation were 1977 through 1983. These years match those used for differentiation when it was a dependent variable regressed on organizational size for the same (1977-83) or earlier years (1976-82 or 1975-81) (Equations 3-5 in the Methods section). The majority (69.4%) of organizations grew in differentiation and the remainder declined. Slightly less than 50 percent of the organizations with growing differentiation had the expected positive relationship between differentiation and administration. In most of these cases, administration grew at a rate proportionately less than the rate at which differentiation grew. The majority of organizations with declining differentiation added proportionately more administrators in spite of decline in number of departments.

Between-Organization Analyses

For comparison with the within-organization analyses reported in the previous section and with previous between-organization research, data from two of the years covered by the within-organization analyses were used to compute regressions representing relationships posited by the theory of structural differentiation. We conducted separate cross-sectional tests of the theory with 1977 and 1983 data. These same years also provided the data for a test using two points of longitudinal data with a between-organization design. As noted in the Methods section, we used logarithmic transformations, absolute values, and ratio measures of administration to match both previous research and our within-organization analyses. Replication of both methods and measurement of earlier between-organization studies was necessary for at least one major reason: to the extent that this study's between-organization findings are consistent with similar previous studies, they suggest that the data examined here are not unique. If they are not unique, a stronger argument can be made when within-organization analyses produce different results than between-organization analyses that cross-sectional comparisons among organizations do not reflect organizational change.

Cross-sectional analyses. Cross-sectional models supported Blau's theory to a greater degree than within-organization models based on the same data, as Table 3 shows. In two separate years, cross-sectional differences in size had a strong positive relationship with differences in differentiation and with differences in absolute sizes of administrative staffs. Unstandardized slopes for the equations in which size and administration were transformed logarithmically, shown in the column headed *LNADMIN*, demonstrated that proportional differences in size produced proportionately smaller administrative staffs. Although positive, the slopes for *LNADMIN* were less than 1. The negative effects in the two equations of administrative ratios (the ratio of administration size to organizational size) on organizational size also confirmed these findings. The regressions of administration size on the absolute amount of differentiation showed that organizations with more differentiation had more administrators even with size controlled. In addition, there were direct effects of differentiation on the administrative ratio. However, equations with logarithmic transformations indicated that the relationships

TABLE 3
Cross-Sectional Equations for Selected Years[a]

		Dependent Variables[b]			
Independent Variables	Years	*DIFF*	*ADMIN*	*LNADMIN*	% *ADMIN*
SIZE[c]	1977	.70	.27	.71 [.42]	-.39
	1983	.73	.33	.61 [.36]	-.35
DIFF[d]	1977	—	.47	.02 [.02]	-.18
	1983	—	.33	.09 [.07]	-.21

[a] Standardized slopes; unstandardized slopes are reported for transformed variables in brackets.

[b] Regressions with administration as the dependent variable included both size and differentiation as independent variables.

[c] Transformed to base 10 log for the regression of *LNADMIN* on size.

[d] Transformed to base 10 log for the regression of *LNADMIN* on differentiation.

between the magnitudes of differentiation and administration were not proportional.

Between-organization analyses with longitudinal data. Table 4 reports eight equations representing two approaches used to study the major relationships posited by the theory of structural differentiation. The analysis employed a between-organization design with the 1977 and 1983 data. In the first approach, variables were transformed logarithmically. These transformed variables bear the prefix *LN*. The gain scores (e. g., $\mathit{LNSIZE}_{83} - \mathit{LNSIZE}_{77}$) represent proportional differences in the between-organization equations. In the second approach, we used absolute numbers for simple gain scores (e. g., $\mathit{SIZE}_{83} - \mathit{SIZE}_{77}$). As did previous research using gain scores or proportional change scores (Dewar & Hage, 1978), this study controlled for the initial level of the change variable. In addition, for both approaches the dependent variable was lagged so that the only potential variance to be explained was the change in the dependent variable that occurred between 1977 and 1983 (cf. Werts & Linn, 1970).

Naturally, the lagged dependent variable was the largest predictor in all equations. Otherwise, neither the levels nor the changes in the independent variables had appreciable effects on the changes in differentiation or administration.

An examination of the distribution of slopes (Table 2) gives some indication of possible reasons for the lack of effects produced by the change variables (gain scores) in the between-organization longitudinal data equations

TABLE 4
Between-Organization Equations for Growing and Declining Organizations[a]

(a) Growing Organizations

(1) $\mathit{LNDIFF}_{83} = .91[.93]\ \mathit{LNDIFF}_{77} + .09[.07]\ \mathit{LNSIZE}_{77} - .00[.02]\ (\mathit{LNSIZE}_{83} - \mathit{LNSIZE}_{77})$

(2) $\mathit{DIFF}_{83} = .80\ \mathit{DIFF}_{77} + .18\ \mathit{SIZE}_{77} + .05\ (\mathit{SIZE}_{83} - \mathit{SIZE}_{77})$

(3) $\mathit{LNADMIN}_{83} = .83[.83]\ \mathit{LNADMIN}_{77} + .05[.03]\ \mathit{LNSIZE}_{77} + .03[.07]\ (\mathit{LNSIZE}_{83} - \mathit{LNSIZE}_{77}) + .03[.02]\ \mathit{LNDIFF}_{77} + .04[.11]\ (\mathit{LNDIFF}_{83} - \mathit{LNDIFF}_{77})$

(4) $\mathit{ADMIN}_{83} = .84\ \mathit{ADMIN}_{77} + .12\ \mathit{SIZE}_{77} + .06\ (\mathit{SIZE}_{83} - \mathit{SIZE}_{77}) - .07\ \mathit{DIFF}_{77} - .02\ (\mathit{DIFF}_{83} - \mathit{DIFF}_{77})$

(b) Declining Organizations

(5) $\mathit{LNDIFF}_{83} = 1.01[1.09]^{b}\ \mathit{LNDIFF}_{77} - .07[.05]\ \mathit{LNSIZE}_{77} + .08[.36]\ (\mathit{LNSIZE}_{83} - \mathit{LNSIZE}_{77})$

(6) $\mathit{DIFF}_{83} = .98\ \mathit{DIFF}_{77} + .02\ \mathit{SIZE}_{77} + .06\ (\mathit{SIZE}_{83} - \mathit{SIZE}_{77})$

(7) $\mathit{LNADMIN}_{83} = .81[.70]\ \mathit{LNADMIN}_{77} + .01[.01]\ \mathit{LNSIZE}_{77} - .11[.32]\ (\mathit{LNSIZE}_{83} - \mathit{LNSIZE}_{77}) + .04[.03]\ \mathit{LNDIFF}_{77} + .13[.30]\ (\mathit{LNDIFF}_{83} - \mathit{LNDIFF}_{77})$

(8) $\mathit{ADMIN}_{83} = .70\ \mathit{ADMIN}_{77} + .12\ \mathit{SIZE}_{77} - .09\ (\mathit{SIZE}_{83} - \mathit{SIZE}_{77}) + .01\ \mathit{DIFF}_{77} + .01\ \mathit{DIFF}_{77} + .09\ (\mathit{DIFF}_{83} - \mathit{DIFF}_{77})$

[a] Metric (unstandardized) slopes are shown in brackets. All other slopes are standardized.

[b] The predicted slope fell out of the range of the data largely due to the presence of an outlier and not because of a lack of variance in differentiation. Over 94 percent of the organizations had variance in differentiation over time. The mean change in differentiation between 1977 and 1983 for declining organizations was loss of 2.3 departments.

(Table 4). Table 2 shows that many organizations had relationships that produced regression slopes in directions hypothesized by Blau's theory. Yet there were almost always as many, if not more organizations, that had opposite effects. These results thus cancelled each other out in the overall estimates of change produced by the between-organization longitudinal data. Cross-sectional models that summarize estimates of change for a sample of organizations seem likely to show little or no effects when subsets of a sample show opposite processes of change.

Conditional Relationships in Blau's Theory

The divergent types of change shown by the within-organization analyses suggested that a search for moderating effects (conditional relationships) might be useful to specify further the conditions under which the theory of structural differentiation had more or less support. Exploratory analyses investigated the theory's basic relationships for growing and declining organizations that were (1) under either public or private control, (2) large or small, and (3) rapidly or slowly growing. In general, results did not show moderating effects except to suggest somewhat greater support for Blau's theory under conditions of rapid growth. The sample size was not big enough to analyze the distribution of slopes across these classifications, however.

DISCUSSION

In general, cross-sectional relationships were consistent with previous research and supported the theory of structural differentiation. Longitudinal analyses failed to support the theory. A detailed consideration of the within-organization findings suggests some possible reasons for this lack of convergence.

The major advantage of studying change using within-organization models rather than between-organization models is that the within-organization models show the processes of structural change in detail. This level of detail allowed us to examine hypotheses derived from the theory of structural differentiation concerning growth and decline. These hypotheses were not stated directly in the original theory, but are implicit if Blau's theory applies to change in general. The advantage of analyzing subtleties of relationships became particularly apparent when we considered the forms of relationships with growth and decline in both size and differentiation.

A reduction in differentiation was expected with a decline in size. This occurred in over 60 percent of the universities. However, in a few it happened proportionately slower than their declines, and some even added subunits during reductions in size. We expected administrative diseconomies with declines in size. This occurred for some universities (39.5%) through either greater or less-than-proportional reductions in administrative staff. However, many more added administrators during declines, perhaps to reform their organizations and stem the loss of students. Assuming that less differentiation would reduce the need for coordination and control and hence reduce the need for administrative staff, we anticipated that declining differentiation would lead to decline in administration size. Again contrary to our expectations, over 50 percent of the organizations added administrators when their levels of differentiation were reduced. Essentially, Blau's theory was not a good predictor of the changes that occurred among structural characteristics during declines of size for these organizations.

For universities growing in size, neither differentiation nor size of administration behaved in line with Blau's theory. Only a few increased differentiation and administrative size by less-than-proportional amounts. Many actually eliminated departments and administrators as a reaction to increases in size. Growth in differentiation, however, did lead to the expected increase in administrative staff for a majority of the organizations studied.

The within-organization findings for administration also gave some indication about how cross-sectional differences may appear through processes unanticipated by Blau. Cross-sectional analyses have often found negative relationships between organizational size and relative size of administrative staffs. The theory of structural differentiation explains this finding as an economy of scale produced by less-than-proportional additions of administrators in growing organizations. However, within-organization analyses suggested that these differences arise over time through processes not considered by the theory of structural differentiation. The organizations studied here often eliminated administrators during growth and added them during decline. Both processes will result in larger organizations having proportionately smaller administrative staffs, even though this result need not come about through the less-than-proportional administrative growth that is a major theoretical point of Blau's theory. This illustrates at least one situation in which the results predicted by the theory can exist even when its explanation of processes is incorrect. We observed somewhat similar processes for the relationships of size and differentation.

The findings for the within-organization analyses also raised an issue ignored by Blau's theory. Using three time lags for organizational size, we found no dominant lag structure for the regressions of size of administration and differentiation on organizational size. It appeared that organizations responded with varying degrees of speed to structural change. Distributed lag models with several lags of size and differentiation included in the model simultaneously (Kmenta, 1971) may clarify patterns of response. It is also possible that lag structures for individual organizations are moderated by factors not considered by the theory of structural differentiation.

Although these findings were derived from a homogeneous population, additional analyses not reported indicated that the findings are generalizable to all colleges and universities. An earlier phase of this research project—using a random sample of colleges and universities, a more inclusive measure of administrative size, and enrollment as the measure of organization size—produced similar findings.

CONCLUSIONS

A comparison of the findings for the cross-sectional analyses and the between- and within-organization longitudinal models suggests that the theory of structural differentiation offers better explanations of scale than of change. That is, the theory better explains differences in scale among organizations than it explains how individual organizations change. Cross-sectional regression equations showed strong support for most aspects of Blau's theory, but both between- and within-organization longitudinal models suggested little support for it. Additionally, although within-organization analyses demonstrated that the vast majority of the organizations did change over time, the patterns of change diverged from those predicted by Blau's theory. In essence, we conclude that these data fail to support the theory of structural differentiation as a theory of change. They suggest instead that it be considered a theory of scale.

One explanation of why Blau's theory can account for differences among organizations but not for changes within organizations may lie within the theory itself. Several of its propositions state that increasing size has a declining effect on other variables—for example, on the administrative economy of scale. Given these propositions, it would seem that most organizations would reach a point where the declining effects of size would become negligible after a sufficient period of growth. As a result, differences in scale would remain among organizations after they grew to some absolute level of size, but future change would have little effect. In a sense, there would exist a threshold; at a given level of size, organizations would develop certain characteristics. Under such conditions, cross-sectional tests would confirm Blau's theory if most organizations observed had histories of growth—a likely situation for most cross-sectional studies because organizations with histories of decline are not likely to have survived at the same rates as growing organizations. Population ecologists might also see other factors as determining cross-sectional differences among organizations. Such factors as norms of organizational structuring at the time of founding and size at birth would account for differences among organizations that do not result from later organizational change. Thus, we suspect that the temporal wording of Blau's theory is misleading. It appears that reaching a certain magnitude or level of size, not changes per se, leads to marginal effects on differentiation and administrative economies. The data comparing the three basic approaches to the study of the theory of structural differentiation seem to support this contention.

The theory implies rational organizational behavior by managers responding to growth or to differences in scale among organizations: Increasing differentiation in response to increasing size presumably takes advantage of the benefits of specialization. More administrators are then necessary to supervise more subunits and to manage the increasingly complex integration of the units. Differentiation aside, increased size alone increases problems of coordination and control and in turn produces a need for a bigger administrative staff. However, once basic administrative functions are in place, organizations only need to add staff marginally and thus achieve administrative economies of scale.

Although our findings indicated that Blau's theory did not explain the relationships observed, these relationships may have resulted from decision making equally as rational as that theory presumes. Adding administrators during a decline phase could be an attempt to offset decline with more administrative expertise. Declining universities might for example need more effort put into student recruitment and alumni solicitation. In addition, reduced departmentalization and resultant elimination of some department chairpeople necessitate new high-level administrators to perform administrative tasks formerly done within departments. Through reducing administrative overhead and numbers of departments, many growing organizations in our research sample were even more economical than anticipated. This high level of economy may be a reaction to the environmental changes of the 1970s. Inflationary pressures and the knowledge that many colleges and universities were declining might have inspired organizational decision makers to shape lean organizations ready to deal with actual and potential losses of resources. However, although changes in size alone do not increase organizational fragmentation or need for administrators, the logistics of managing more subunits does have its expected effect on need for administrators.

Rational explanations for why the majority of organizations studied did not follow the processes of change implied by Blau's theory are essentially hypotheses of moderating or contingent relationships. Our results bring to mind earlier studies challenging the appropriateness of one model of bureaucracy, such as the Weberian, for all situations (e. g., Stinchombe, 1959). Rational reactions, in terms of administrative needs, to changing organizational size and differentiation may be consistent with the theory of structural differentiation under some conditions, but equally rational reactions may fail to support the theory under other conditions. Our brief exploration of moderating effects suggested that rapid growth leads organizations to behave in a manner congruent with Blau's theory. These organizations may be approaching the levels of size and differentiation that produce the associations with each other and with size of administration observed with cross-sectional data and hypothesized by his theory.

Of course, rational behavior does not completely govern organizations, nor do organizational decision makers accurately perceive their environments in all conditions (Astley & Van de Ven, 1983). Misperceptions of an environment (Rubin, 1979), historical traditions (Clark, 1972), and internal and external political behavior (Pfeffer & Salancik, 1978) are all possible moderating or intervening variables that could alter the relationships hypothesized by Blau for size, differentiation, and administration. In organizations that are in decline or experiencing scarcity of resources, power rather than reason may have more influence on decisions regarding retention of administrators or the elimination of departments (Ford, 1980b; Hannan & Freeman, 1978; Pfeffer & Salancik, 1974).

Although the data examined here did not support Blau's hypotheses when applied to changes within organizations, future replications of this research will be necessary for a more conclusive refutation of the theory of structural differentiation as a theory of change. The general lack of strong support for the theory shown by the within-organization analyses may indicate that longer lag periods should be investigated. However, as a theory derived from cross-sectional data, Blau's theory fails to suggest appropriate lag structures. Thus, for future tests of this theory as a theory of change, researchers and theorists must specify the time frame in which they expect effects of changing size and differentiation to become manifest. Without such specification, it can always be claimed that an insufficient lag period was observed. This study suggests that for many organizations any effects of changes in size may occur, if at all, over the course of more than two years.

This investigation was limited in focus and generalizability. We studied only the three major relationships hypothesized by the theory of structural differentiation and used only one type of organization. The within-organization analyses did not consider the theory's hypotheses concerning span of control, homogeneity and average size of subunits, vertical differentiation, and feedback effects. As such, future within-organization research is needed to assess more fully the theory's implications for change. In addition, since the results of this study pertain only to universities, future researchers should consider using other types of organizations. Such investigators might give special concern to vertical differentiation because academic organizations have limited variance on this variable. Attention might also be given to identifying organizational types with varied career options. Managers and administrators of nonacademic organizations, not having the typical academics' option of returning to tenured teaching positions, might well behave unlike the academic administrators observed in this study—especially during decline.

Finally, in more general terms, it should be noted that the lack of congruence between cross-sectional and longitudinal findings for Blau's theory does not indicate that one type of research design is preferable, nor does it indicate that theories of scale are less worthy of concern. Rather, these findings highlight the need to develop theories of organizational structure that consider both magnitude of scale and degree of change in the variables of interest. Empirical tests of such theories will necessarily make comparisons between organizations as well as observe changes within organizations.

REFERENCES

Akers, R., & Campbell, F. L. 1970. Size and the administrative component in occupational associations. ***Pacific Sociological Review***, 13: 241–251.

Assmus, G., Farley, J. U., & Lehmann, D. L. 1984. How advertising affects sales: Meta-analysis of econometric results. ***Journal of Marketing Research***, 21: 65–74.

Astley, W. G., & Van de Ven, A. H. 1983. Central perspectives and debates in organization theory. ***Administrative Science Quarterly***, 28: 245–273.

Bedeian, A. G. 1980. ***Organizations: Theory and analysis.*** Hinsdale, Ill.: Dryden.

Beyer, J. M., & Trice, H. M. 1979. A reexamination of the relations between scope and various components of organizational complexity. ***Administrative Science Quarterly,*** 24: 48–64.

Blau, P. M. 1970. A formal theory of differentiation in organizations. ***American Sociological Review,*** 35: 201–218.

Blau, P. M. 1972. Interdependence and hierarchy in organizations. ***Social Science Research,*** 1: 323–349.

Blau, P. M. 1973. ***The organization of academic work.*** New York: John Wiley & Sons.

Blau, P. M. 1983. Citation classic — A formal theory of differentiation in organizations. ***Current Contents/Social and Behavioral Sciences,*** 19: 20.

Blau, P. M., & Schoenherr, R. 1971. ***The structure of organizations.*** New York: Basic Books.

Clark, B. R. 1972. The organizational saga in higher education. ***Administrative Science Quarterly,*** 17: 178–184.

Cullen, J. B., & Baker, D. D. 1984. Administration size and organization size: An examination of the lag structure. ***Academy of Management Journal,*** 27: 644–653.

Daft, R. L., & Becker, S. W. 1980. Managerial, institutional, and technical influences on administration: A longitudinal analysis. ***Social Forces,*** 59: 392–413.

Dewar, R., & Hage, J. 1978. Size, technology, complexity, and structural differentiation: Toward a theoretical synthesis. ***Administrative Science Quarterly,*** 23: 111–136.

Ford, J. D. 1980a. The administrative component in growing and declining organizations: A longitudinal analysis. ***Academy of Management Journal,*** 23: 615–630.

Ford, J. D. 1980b. The occurrence of structural hysteresis in declining organizations. ***Academy of Management Review,*** 5: 589–598.

Freeman, J. 1979. Going to the well: School district administrative intensity and environmental constraint. ***Administrative Science Quarterly,*** 24: 119–133.

Freeman, J. H., & Hannan, M. T. 1975. Growth and decline processes in organizations. ***American Sociological Review,*** 40: 215–228.

Freeman, J. H., & Kronenfeld, J. E. 1973. Problems of definitional dependency: The case of administrative intensity. ***Social Forces,*** 52: 108–121.

Furness, W. T. (Ed.) 1973. ***American universities and colleges*** (11th ed.). Washington, D. C.: American Council on Education.

Gillespie, D. F., & Mileti, D. S. 1976. A refined model of differentiation in organizations. ***Sociology and Social Research,*** 60: 263–278.

Goldman, P. 1973. Size and differentiation in organizations: A test of a theory. ***Pacific Sociological Review,*** 16: 89–105.

Hannan, M. T., & Freeman, J. H. 1978. Internal politics of growth and decline. In M. W. Meyer & Associates (Eds.), ***Environments and organizations:*** 177– 199. San Francisco: Jossey-Bass.

Hendershot, G. E., & James, T. F. 1972. Size and growth as determinants of administrative-production ratios in organizations. ***American Sociological Review,*** 37: 149–153.

Institute for Scientific Information. 1984. ***Social science citation index 1983 annual.*** Philadelphia: Institute for Scientific Information.

Jackson, J. H., & Morgan, C. P. 1982. ***Organization theory.*** New York: Prentice-Hall.

Kimberly, J. R. 1976. Organizational size and the structuralist perspective: A review, critique, and proposal. ***Administrative Science Quarterly,*** 21: 571–597.

Kmenta, J. 1971. ***Elements of econometrics.*** New York: Macmillan Book Publishing Co.

MacMillan, A., & Daft, R. L. 1979. Administrative intensity and ratio variables: The case against definitional dependency. ***Social Forces***, 58: 228–248.

Marquis Academic Media. 1975/76–1983/84. ***Yearbook of higher education***. Chicago: Marquis Who's Who.

Meyer, M. W. 1971. Some constraints in analyzing data on organizational structures: A comment on Blau's paper. ***American Sociological Review***, 36: 294–297.

Meyer, M. W. 1972. Size and the structure of organizations: A causal analysis. ***American Sociological Review***, 37: 434–441.

Mileti D. S., Gillespie, D. F., & Haas J. E. 1977. Size and structure in organizations. ***Social Forces***, 56: 208–216.

Miller, G. A., & Conaty, J. 1980. Differentiation in organizations: Replication and cumulation. ***Social Forces***, 59: 265–274.

Miller, D., & Friesen, P. H. 1982. The longitudinal analysis of organizations: A methodological perspective. ***Management Science***, 28: 1013–1034.

Ostrom, C. W. 1978. ***Time series analysis: Regression techniques***. Beverly Hills, Calif.: Sage.

Pfeffer, J., & Salancik, G. R. 1974. Organizational decision making as a political process: The case of a university budget. ***Administrative Science Quarterly***, 19: 135–151.

Pfeffer, J., & Salancik, G. R. 1978. ***The external control of organizations***. New York: Harper and Row.

Rubin, I. S. 1979. Retrenchment, loose structure and adaptability in the university. ***Sociology of Education***, 52: 211–222.

Rushing, W. A. 1966. Organizational size and administration: The problems of causal homogeneity and a heterogeneous category. ***Pacific Sociological Review***, 9: 100–108.

Stinchombe, A. L. 1959. Bureaucratic and craft administration of production: A comparative study. ***Administrative Science Quarterly***, 4: 168–187.

Tsouderos, J. E. 1955. Organizational change in terms of a series of selected variables. ***American Sociological Review***, 20: 206–210.

Werts, C. E., & Linn, R. L. 1970. A general linear model for studying growth. ***Psychological Bulletin***, 73: 17–22.

John B. Cullen is an associate professor of management at the University of Nebraska-Lincoln. The dynamics of structural change in organizations are the foci of his recent research. He received his Ph. D. degree from Columbia University in 1977.

Kenneth S. Anderson is a Ph. D. candidate in management at the University of Nebraska-Lincoln and is an assistant professor of management at Gonzaga University. His research interests include the processes of change in organizations, the influence of structure on employee turnover, job design, and organizational stress.

Douglas D. Baker is an assistant professor of management at Washington State University. He received his Ph. D. degree from the University of Nebraska-Lincoln in 1983. His current research focuses on organizational life cycles, individual motivation, and sexual harassment in organizations.

© *Academy of Management Journal*
1986, Vol. 29, No. 2, 230–261.

TACTICS OF IMPLEMENTATION

PAUL C. NUTT
Ohio State University

This study identified implementation tactics used by managers in making planned changes by profiling 91 case studies. Analysis revealed four types of tactics, variations within these tactics, and hybrid tactics. These tactics are described and related to their success rates and conditions surrounding their use. The study found a 100 percent success rate when key executives used an intervention tactic, but observed this tactic in less than 20 percent of the cases. Both the persuasion and participation tactics had 75 percent success rates; persuasion had the highest frequency of use, 42 percent, and participation the lowest, 17 percent. Both made high demands on resources. Implementation by edicts had a 43 percent success rate and a 23 percent frequency of use. Intervention tactics and their variations were effective for all types of changes and under varying levels of time pressure and importance, suggesting that managers should use these tactics more often.

Managers, responding to needs and opportunities, revitalize organizations by setting in place planned changes that alter their organizations' products, services, internal operations, and policies. Researchers have given the process of planned change various labels: planning (Churchman, 1979a), organizational change and regeneration (Lewin, 1947), and decision making (March, 1981). Changes become the proposals for action that managers implement. Implementation is a series of steps taken by responsible organizational agents in planned change processes to elicit compliance needed to install changes.

Previous research has reported many implementation failures (Bardach, 1977; Pressman & Wildavsky, 1973; Schultz & Slevin, 1975; Zand & Sorensen, 1975). These failures can be attributed (1) to individuals or groups who attempt to maintain relationships, sustain existing procedures, or retain control over valued activities that would be altered, and (2) to misunderstandings or disagreements about expected benefits (Greiner, 1970; Kotter, 1979; March, 1981). People in organizations use gesture, delay, and obstruction to contain or block change attempts they find threatening or merely disagreeable (Bardach, 1979). To be successful, managers must devise tactics that neutralize or at least contain people who delay making essential commitments, protect turf, posture, or carry out vendettas.

This paper reports on a long-term research effort that documented the steps or tactics managers take to implement planned changes. The purpose of

---

The author thanks the referees for their many helpful comments on previous versions of this paper.

this research was (1) to identify and describe the implementation tactics managers use, (2) to determine the relative success of these tactics, and (3) to identify conditions under which various tactics are used. Profiles of 91 case studies of planned change provided data for this research. I identified implementation activities in the cases, classified them according to implementation tactics that had distinct procedures, and determined each tactic's success rate. Finally, I explored the relationships between tactics and contextual factors to determine if type of change, time pressure, quality of staff support, perceived importance, and size of budget for change processes led to the use of particular tactics.

THE STUDY OF IMPLEMENTATION

Implementation research has devised prescriptions based on logic, developed administrative tools, studied factors that influence adoption, and explored how changes are put into effect.

The prescriptive approach (e.g. Bennis, 1973; Kotter & Schlesinger, 1979) advocates the use of particular mechanisms to make changes. For example, Bourgeois and Brodwin (1984) presented commander, change, collaborator, cultural, and coercive models as alternative implementation approaches for executives. Nutt (1983) offered a contingency model that defined sixteen environments in terms of differences in climate and fitted implementation approaches described in the literature to each environment. The prescriptive literature has offered many useful ideas for implementation, but these ideas are based on logic rather than on data that supports the use of specific guidelines, recommendations, and assertions.

The tool approach considers the value of techniques, such as t-groups, that deal with resistance to change and other issues related to implementation. For instance, Schein's (1961) studies of brainwashing during the Korean conflict demonstrated the value of breaking old social ties and establishing new ones in conversion attempts. Daft and Becker identified the roles of incentives and other factors that they call "enablers" (1978:8) in promoting implementation success. Such studies have identified useful ways for managers to support implementation processes. However, the tool approach creates the impression that implementation is nothing more than the selection and use of departmentalization, reward systems, coordination, and the like.

The factor identification approach surveys managers to determine factors that appear significant and correlates these factors with success and failure. For instance, Zand and Sorenson (1975) developed questionnaires to elicit managers' rankings of factors they believed to be important. Their questionnaire also sought to determine the extent of organizational compliance with changes, and of internalization or adoption of changes. Researchers have often identified factors in such questionnaires by stages in a prescriptive model; for instance, Zand and Sorenson used a model based on the work of Lewin (1947) and Schein (1961). Other studies of this type have built models that depicted steps believed to be essential in change processes and used these models to elicit factors (e.g., Beyer & Trice, 1978). Such studies have found that success seems to be related to a positive balance in managers' feelings between desirable and undesirable factors such as confidence, fear, and understanding. They have also identified the importance of pivotal actors in change processes, and distinguished between intended and actual uses. However, Ginzberg (1975), Lucas (1975), and other reviewers have found little consensus among many well-executed studies as to which of the hundreds of factors that have been identified are crucial to success or precursors to failure.

Others have used case studies to search for implementation approaches (e.g., Argyris, 1970). In fact, much of what we know about implementation comes from field studies of managers attempting to make changes. The ways managers carry out implementations reveal their procedures. The case study approach assumes that managers learn to be effective through "practice heuristics" (Schon, 1983:68) that evolve procedures to deal with implementation and other issues. However, interviews with managers have failed to surface explicit procedures (Keen & Scott-Morton, 1978), forcing researchers to use indirect methods, such as retrospective reconstruction of events. Investigators have studied successful and unsuccessful implementation attempts to try to derive the procedures used. For example, Pressman and Wildavsky (1973) followed the implementation of legislation to detect obstacles that were erected and the rationales behind these obstacles in order to suggest ways to overcome these barriers and enhance prospects for successful implementation. Bardach (1979) construed implementation as a game and described steps that promote winning. Many interesting concepts have emerged from this type of study. However, as Keen and Scott-Morton pointed out, the idiosyncractic nature of case studies limits researchers' abilities to pin down procedures underlying successful implementations. A much larger data base is required to understand the finer points of implementation and distill steps that can be unequivocally linked to success. Because of these limitations, implementation prescriptions offered by Greiner (1970) and others often take the form of reversing the steps used in failed change attempts.

Implementation research has provided valuable guidance and identified pitfalls and opportunities, but has provided few insights into how managers carry out implementation. To extend implementation research, three changes in approach seem essential. First, the relationship between developmental and compliance-enhancing activities must be considered by researchers exploring formulation and goal setting, environmental assessment, option development and selection, evaluation, and other stages of change-making processes. Many process-based studies have failed to capture this relationship, thereby ignoring potentially crucial implementation steps that managers carry out during development. Second, investigators should frame change processes broadly so implementation activities can be seen as part of overall processes. The backdrop of a change process allows implementation steps to stand out in sharp relief, making it relatively easy to identify and classify activities undertaken by managers to elicit the support, cooperation, or acquiescence required to install changes. Third, past studies of change failed to recognize that, although change processes may continue working their way through organizations, particular managers' involvement with them has beginnings and ends. Fourth, some of the practices often used in case study research have serious limitations. Several studies (Bower, 1970; Mintzberg, Raisinghani, & Theoret, 1976; Soelberg, 1967; Witte, 1972) popularized a technique that immerses a researcher into the raw data describing each case. This type of research allows patterns to emerge by examining and comparing each case to the other cases. This approach is infeasible for large numbers of cases owing to the mass of detail that must be considered. The alternative of limiting data bases to make them manageable produces conclusions of suspect generalizability. Researchers using data bases with few cases may fail to discover important implementation-related steps or fail to recognize the idiosyncratic nature of the steps they do discover. Alternatively, reseachers can impose a framework to search for implementation steps instead of allowing patterns to emerge. However, imposing a framework may result in losing important messages that do not fit classification schema. In this research the advantages of systematic description and of using a large data base were thought to outweigh the possibility of lost information.

THEORETICAL FRAMEWORK

Implementation

Managers use implementation to make planned changes in organizations by creating environments in which changes can survive and take root. Planned changes are proposals for action that alter organizations' practices. A change can be adaptive, calling for installation of known practices (Beyer & Trice, 1978); innovative, introducing practices that are new to adopting organizations (Damanpour & Evan, 1984); or radically innovative, introducing practices new to all organizations in the same business or industry (Zaltman, Duncan, & Holbek, 1973). Changes can involve (1) technical innovations resulting from the application and use of technology; (2) administrative innovations dealing with relationships and promulgating new rules, roles, procedures, or structures within organizations; or (3) adaptation of technology or managerial practices used by competing or cooperating organizations.

Previous research has defined implementation in terms of actors, intents, and procedures (Ansoff, 1984; Beyer & Trice, 1978; Daft & Becker, 1978; Zaltman et al., 1973). Implementation is a procedure directed by a manager to install planned change in an organization. There is widespread agreement that managers are the key process actors and that the intent of implementation is to install planned changes, whether they be novel or routine. However, procedural steps in implementation have been difficult to specify because implementation is ubiquitous. Implementation has figured prominently as part of planned change processes labeled decision making (Mintzberg et al., 1976), strategic management (Schendel & Hofer, 1979), organizational regeneration (March, 1981), change (Zand & Sorensen, 1975), legislation and policy initiatives (Bardach, 1977), and diffusion of innovation (Waters, Salipante, & Notz, 1978). Ansoff made several important distinctions pertinent to these processes of planned change, identifying four procedures, called the entrepreneurial, exploration, control, and implementation subprocesses. From this perspective, implementation can be viewed as a procedure used in a planned change process that lays out steps taken to entice stakeholders to support changes. A coherent set of steps becomes a tactic used by managers to elicit the support, cooperation, or acquiescence needed to insure compliance with planned changes. Tactics should capture the arranging and manuevering steps that managers take to deal with social and political issues provoked by change attempts. In this study emphasis is on identifying and describing tactics and not on measuring behaviors and discrete events such as numbers of meetings, analysis of minutes, timing of change processes, or jockeying of stakeholders.

A Transactional Path Model

To determine how managers who sponsor changes promote compliance, this study examined how responsible agents regulate and control a process of planned change. Because steps taken to promote compliance can be found in any stage of a change process, this study treated each change process as a window through which implementation tactics can be viewed.

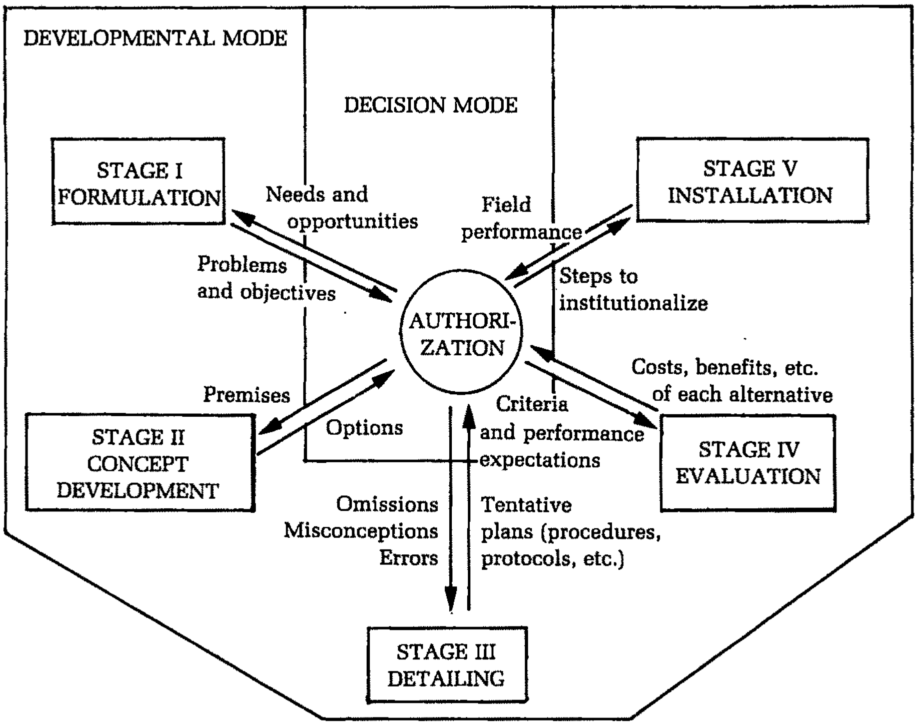
The framework used to describe the planned change process was developed by synthesizing stages from related processes (Nutt, 1984a). The recommended staging of activities was derived from the literature on innovation (Zaltman et al., 1973), policy formation (Dunn, 1981), organizational change (Hage & Aiken, 1970), planning (Nutt, 1984a), and decision making (March, 1981). The specifications for stages and essential information in each stage were distilled for these processes.[1] Figure 1 shows the planned change process as a series of transactions, instead of the static events often represented in similar research (Mintzberg et al., 1976). The arrows in Figure 1 show the direction of information flow between a sponsor-manager in the decision mode and a support team[2] in the developmental mode. Development can involve assessing environments to find niches ripe for exploitation; analyzing causal factors lying behind poor performance; offering ways to restructure products or services that overcome performance problems; making lists of equipment or raw material vendors that meet preset criteria; providing assessments of potentially viable options like vendors, sites, programs, and so forth. This transactional representation has an action-taking character that

---

[1] Readers interested in a detailed discussion of how these stages map to the stages used by others are referred to Nutt (1984a).

[2] Support teams can include technical staff, other managers, or even sponsors acting as their own technical advisors. The make-up of support teams working in the developmental mode can shift with every process stage, or a fixed relationship, such as sponsor-controller, can be maintained throughout a change process.

FIGURE 1
The Transactional Planned Change Process [a,b]


DEVELOPMENTAL MODE
DECISION MODE
STAGE I FORMULATION
Needs and opportunities
Problems and objectives
Field performance
STAGE V INSTALLATION
Steps to institutionalize
AUTHORI-ZATION
Costs, benefits, etc. of each alternative
Premises
Options
STAGE II CONCEPT DEVELOPMENT
Criteria and performance expectations
STAGE IV EVALUATION
Omissions Misconceptions Errors
Tentative plans (procedures, protocols, etc.)
STAGE III DETAILING

[a] Adapted from Nutt (1979, 1984[a]).

[b] Arrows show direction and type of information flow from authorization, shown as circle, in the decision mode to development stages, shown as boxes, in the developmental mode of the transactional process.

describes how managers intervene during development to promote implementation and deal with other issues. The model is transactional because it identifies information that support teams report when carrying out each stage of a change process. Managers or sponsors receive this information and make decisions as change processes evolve. They take action through these decisions.

Transactions. In stage I, formulation, a support team in the developmental mode clarifies the needs and opportunities a manager stipulates in the decision mode. The support team identifies problems and proposes objectives. In stage II, concept development, a sponsor states his or her premises, which identify ways to deal with the problems or respond to the objectives identified in stage I. The support team responds by offering one or more options, which the sponsor considers. In stage III some of the options are selected to be detailed and offered to the sponsor, who tests them for omissions, misconceptions, and errors. Alternatives that can be modified to overcome the sponsor's objections are subject to evaluation in state IV. The sponsor identifies criteria and the support team applies them to assess the merits of alternatives that survive stage III. In the installation stage the sponsor-manager applies tools such as rewards and incentives, personnel selection and promotion, resource allocation, sanctions, coordination, and delegation to put changes into operation. Stage V begins with an attempt to set in place a preferred option and ends when field performance is judged to be adequate.

Paths. The path that transactions take through the model trace the steps in a change process. There are two types of traces (Nutt, 1984a:442). A normative trace identifies what theorists believe should take place during a change process (Churchman, 1979a); it indicates information exchanges and directionality of the exchanges that theorists consider ideal, as shown in Figure 1. A descriptive trace profiles paths actually taken. Managers acting as sponsors can adopt active or passive roles in process management. A passive role would call for delegation. An active sponsor would participate in each stage of a change process, regulating it by acting on the information provided and offering guidance. Between these extremes lie a variety of management approaches. Patterns of information exchanges describe the degree and type of control used to promote compliance and show how managers deal with implementation-related questions. This study analyzed these patterns to identify implementation tactics.

METHODS

This research used a multiple case study approach to identify implementation tactics because it captures the best features of both case study and factor identification approaches: firsthand knowledge and rich descriptive information, and the ability to quantify and generalize. The change process framework provided a common standard for representing cases. Cases were described using this framework to permit a search for tactics related to implementation. Information describing contexts was also collected. Contextual factors were used to specify the motivations and conditions associated with each case so implementation tactics identified in the study could be related to these motivations and conditions. Lastly, the study linked outcomes— success or failure—with tactics.

Process Reconstruction

The research identified the implementation tactics of managers by reconstructing the nature and sequence of activities found in the cases. This approach is similar to meta-analysis (Glass, McGaw, & Smith, 1982), which attempts to integrate findings from independent studies, using each study as a unit of analysis. In process reconstruction, cases are the units of analysis. However, process reconstruction differs from meta-analysis in that random sampling of cases is not used because such an approach to case selection is not feasible.

Process reconstruction required three key activities. The first was to conduct interviews with people who were intimately involved with the details of each change. These interviews identified critical steps, determined the order in which these steps were carried out, and fitted sequences of steps to a classification framework to represent key transactions in each planned change process. The second step was to isolate activities undertaken to elicit or promote compliance in these transactions. The third was to analyze the content of cases with similar and dissimilar transactional patterns to distill these compliance-related steps into tactics and to break tactics down into categories called variations.

Pattern Identification and Classification

Identification of cases. Cases came from organizations that participated in student intern programs. All were service organizations like hospitals, government or nonprofit agencies, charities, and professional societies, so the generalizability of this study's findings may only extend to such service providers. Because the organizations were widely dispersed across the United States and Canada, a broad range of planned change practice was likely represented.[3]

A chief executive, operating, or financial officer (CEO, COO, or CFO) of each of the 91 organizations was contacted by the author and asked to participate in a study that would profile one of their changes.[4] All agreed to participate. The sanctioning executives were asked to select a change that had considerable importance for their organizations; no other guidelines for case selection were offered. Allowing contact persons to select the cases helped to ensure that participants would be cooperative. When executives offered several options, change attempts that added diversity to the data base, both in terms of case types and outcomes, were selected.

Case-related data collection. Multiple interviews were conducted in each organization to minimize memory distortion and memory failure, the two most common errors in reconstructing events (Bartlett, 1954). Each organizational contact named two people who were intimately familiar with each selected case for in depth interviews. One interviewee was the sponsor, the manager who had been administratively responsible for the change. Separate interviews with the sponsor and one other participating executive were conducted.[5] In the unstructured portion[6] of an interview, the interviewee was asked to spell out the sequence of steps that had been taken. The information gleaned from each interview was then summarized by the author using the transactional path model. A path through the model representing the sequence of events recalled by each executive was sketched to provide a picture of each planned change process. The resulting transactional description was then presented to that executive to verify that it accurately represented the process as he or she recalled it. The transactions and their path were modified until the process representation was acceptable to the informant. No more than a week was allowed to elapse between interviews and verification of corresponding path models.

---

[3] The cases in this study, as in most multiple case studies (e.g., Mintzberg & Waters, 1982), do not represent a random sample. A positive bias may be present because organizations that participate voluntarily are likely to share information about change attempts they believe were high quality.

[4] In the interviews the author asked the executive to identify a project in which a change was attempted.

[5] The author pilot-tested the procedure in face-to-face discussions and then carried it out by telephone.

[6] The next section of this paper discusses the structured portion of the interview.

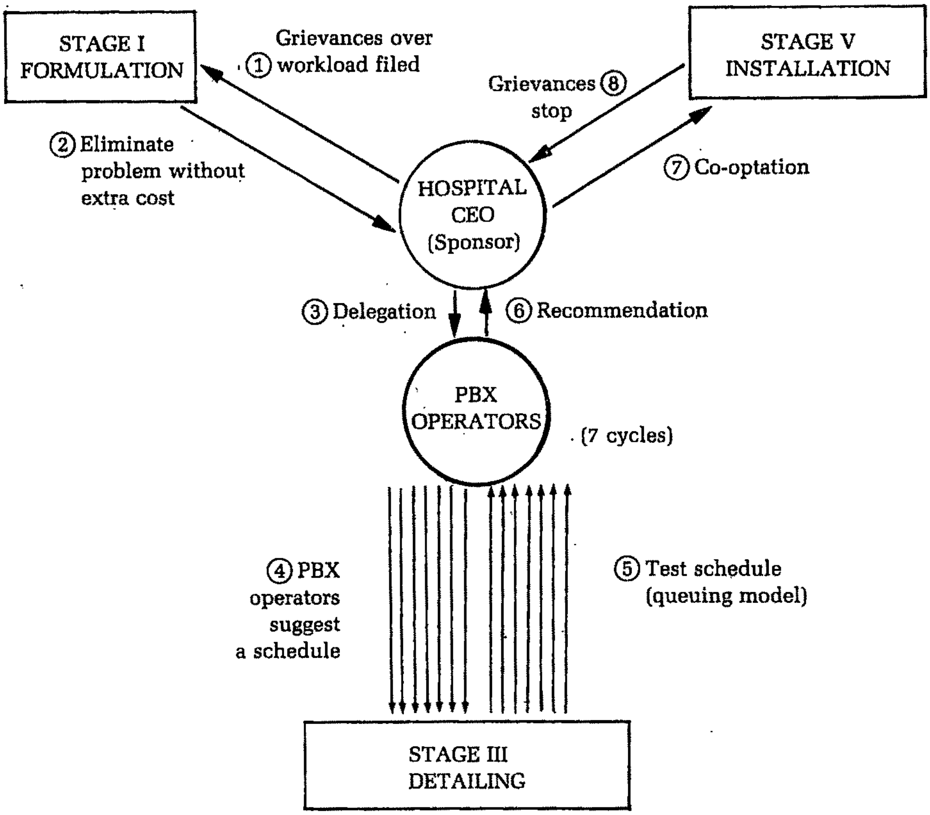
The processes sketched by the two executives were next compared and differences identified. An additional interview reconciled any difference that occurred. In this interview, the executives were asked to modify the transactions until a consensual profile emerged.[7] When differences persisted, others involved in the same change process were interviewed. This interview and other data sources like reports and records were used to determine which version of events seemed the most plausible. If disagreements persisted or if a case lacked detail, that case was to be discarded, but none were. All 91 cases were retained for the classification phase of the study and all met two tests: (1) participants agreed on what had been done, and (2) sufficient detail existed to explain what had been done. Data collection and reduction required about ten hours per case. Data were collected over a period of seven years.

A sample profile. Figure 2 illustrates the type of detail captured in a transactional representation of a planned change with one of the cases in the data base. The direction of the arrows in Figure 2 shows the flow of information through the decision mode. The PBX case began when the organization's telephone operators filed a series of grievances concerning workload (step 1). Discussion with Bell system technical staff revealed that the work demands made on the operators were 20 percent below that expected by Bell for their own operators. With this information in mind, the sponsor, who was the CEO, asked an operations research staff specialist to rectify the situation. The specialist could make any change that did not increase operating costs (step 2). At the specialist's request, the sponsor delegated the choice of the terms of work to the PBX operators, making them a part of the decision mode (step 3). The specialist asked the operators to identify schedules, with the understanding that any schedule could be adopted as long as it did not increase costs. A queuing model was used to evaluate the cost of each proposed schedule in step 5. The seventh option proposed was found to have acceptable levels of both cost and satisfaction; no attempt was made to find the least-cost or optimal schedule. In step 6, this recommendation was presented to the CEO. In step 7, a disgruntled supervisor who had not been involved, reluctantly put into effect the schedule the operators and management

---

[7] When differences emerged, the executives were sent the path models and met to work out their differences.

FIGURE 2
The PBX Case[a]



[a] Arrows trace the flow of information from authorization, designated by circles, to development, designated by boxes.

had agreed on. After the new schedule was implemented the grievances stopped, ending the process in step 8. The manager-sponsor was satisfied when the grievances disappeared and expressed no interest in isolating their cause.

Sorting the profiles. In multiple case study research, reliable classification is crucial (Yin, 1981). In an attempt to identify categories that have classification precision, the 91 profiles or transactional path models were repeatedly sorted until distinct categories of implementation-related activities emerged. To make these sorts, two criteria were applied: (1) convergence and divergence, and (2) the importance of the distinctions (Campbell & Fiske, 1959). The convergence-divergence criterion was met when a classification had explicit similarities within categories and explicit differences among categories. The importance dictum required that the distinctions among categories have practical as well as theoretical importance (McKelvey, 1978).

The transactional path model for each case was reviewed and placed with others that used similar implementation steps. Profiles with ambiguous classifications were reviewed again in an attempt to create new categories or to match existing ones. After each sort, a code number designating a category was placed on the back of each transactional path model to identify it. The sorting process was repeated a year later without reference to the first sort to check intrarater reliability. These classifications matched perfectly. When a second rater sorted the profiles to check interrater reliability, agreement between raters exceeded 90 percent, suggesting that the categories identified were defensible.

Measuring Effectiveness and Conditions of Use

Data collected in the structured portions of the interviews with executives allowed preliminary assessments of the merits of implementation tactics. Descriptions of the contexts of cases determined conditions under which each implementation tactic had been used. In addition, each case was classified as to type and as to success or failure.[8] The data on contextual factors, types of changes, and success rates in Table 2 allowed a preliminary assessment of the effectiveness of tactics under the conditions present when the tactics and variations were used.

A review of each case determined its type; three categories emerged from this analysis: In equipment cases, organizations acquired major pieces of equipment. Construction cases had the primary intent of altering the physical features of key facilities. Program cases provided new services or internal operations and had less than one-quarter of their cost allocated to equipment or construction expenses.

The remaining contextual factors were identified with a survey that participating executives filled out at the end of their interviews. The survey asked respondents to indicate whether change attempts had been adopted or rejected[9] and to specify values for the contextual factors. The survey asked questions using anchored rating scales to determine importance, staff support, time pressure, and process budgets. Executives were asked to estimate a change's importance on a scale ranging from 4 = critical to 1 = routine; the quality of staff support available on a scale ranging from 4 = good to 1 = poor; and the time pressure to complete the change on a scale ranging from 4 = very high to 0 = none. They were also asked to compare the funds allocated to carry out development with funds for other change attempts they had managed. This scale ranged from 5 = far above to 1 = far below.

---

[8] This procedure assumes that a sponsor can define success without using criteria beyond "put into use." The cases bear out this assumption. For example, a new burn center was a failure if the center failed to open; a management information system was not adopted when the organization continued to use its manual system; and a merger was termed a failure when it was not completed.

[9] Measures such as frequency of use and extent to which changes become institutionalized could reveal insights into the degree of adoption. Attempts to collect information for these measures of status proved to be infeasible. For example, when a change in organizational structure was attempted, executives found it impossible to estimate how often the old reporting relationships had been used. To measure frequency of use would have demanded on-site observation, which was infeasible in a retrospective study. Measures of the degree to which objectives were met were also attempted. Initial and final estimates of performance gaps, such as cost or utilization expectations compared to realizations, were collected but proved to have questionable reliability. Records laying out expectations were seldom available making the proportion of objectives realized impossible to gauge. Subjective estimates always moved in favorable directions, even when changes were unsuccessful, making them seem suspect. Disagreements among the executives on these estimates were common and seldom fully resolved.

The estimate-discuss-estimate procedure was used to improve the precision of these estimates (Gustafson, Shukla, Delbecq, & Wallster, 1973). First, executives filled out surveys. After tabulation, discussions of the average values for time pressure, need to act, staff support, and level of process support brought mutual understandings about the processes to the surface. After discussion, executives filled out the surveys a second time. The analysis used values from this second survey.

Analysis

Statistical analyses of the data treated two kinds of tactic variables—tactics and tactical variations, which were subclassifications within tactics—as factors. Levels for each of the tactical types were correlated with the other factors. Analyses related success rate, perceived importance, staff support, time pressure, and process budgets to the tactic variables. Ten one-way ANOVAs were performed, each using one of the two kinds of tactic variables and one of the context or outcome variables. Both parametric and nonparametric analyses followed an ANOVA format, using *F* tests and Kruskal-Wallis chi-square approximations, respectively. Similar significance levels for these tests would suggest that the values for the dependent variables meet interval-scale assumptions. Duncan multiple range tests were used to isolate differences in dependent variables associated with levels of tactic variables.[10]

IMPLEMENTATION TACTICS USED BY MANAGERS

Several distinct types of implementation tactics and within-tactics variations emerged from the analysis. The nature and degree of their sponsors' involvement in process management were found to be distinguishing features of these tactics. Four implementation tactics were used in 93 percent of the cases. These tactics were called intervention, participation, persuasion, and edict. Each tactic had important variations. Table 1 lists the frequency of use for each tactic and tactical variation and summarizes their salient features. The next subsection lays out the steps used in each of these implementation tactics and their variations. Cases drawn from the data base illustrate their distinctive features.[11]

---

[10] In a Duncan multiple range test (DMRT) the equivalent of multiple *t* tests are conducted for each pair of the levels of the independent variable to find tactics that have similar and dissimilar values for each dependent variable. An .05 level of significance is used. The DMRT controls for the increased probability of generating significant findings by running multiple comparisons.

Implementation by Intervention

The tactic. Implementation-related steps were found to begin and end in stage V for this tactic (Figure 3). Issues surrounding compliance, such as the arena of inquiry and ideas that must be considered to win over stakeholders, constrained activity in the developmental stages. In these cases, manager-sponsors were careful to acquire sanctions to activate and regulate the process, and to make it clear to all participants, such as other managers, who had control.

To initiate change processes, the manager-sponsors became protagonists by creating rationales for action in the minds of key people. Appraising performance levels (Figure 3, step 1) and demonstrating performance inadequacies

TABLE 1
Features of Implementation Tactics

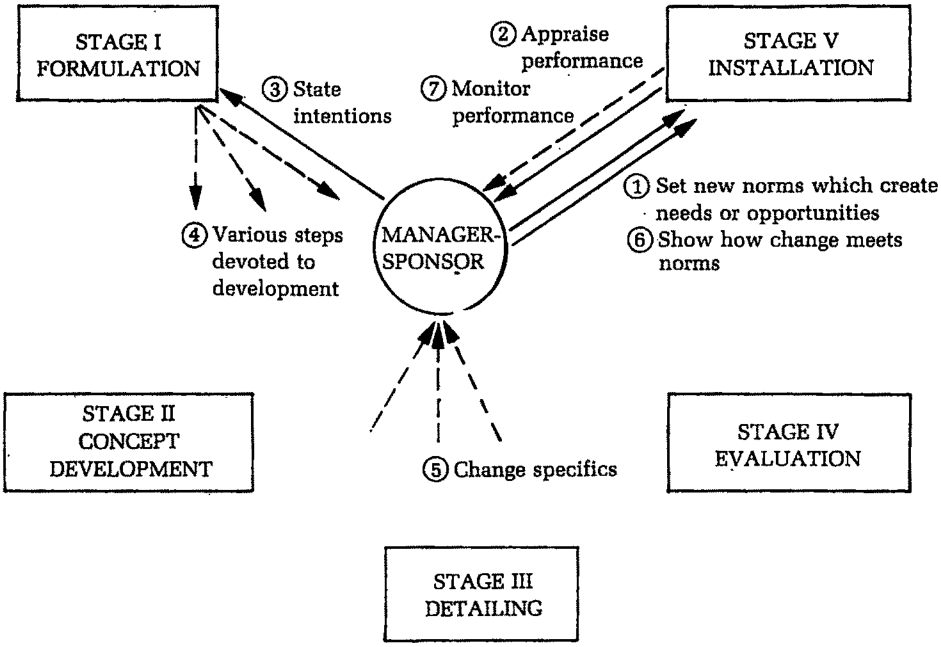
Tactics and Variations	Frequency of Occurrence[a]	Salient Features
Implementation by intervention	19	Key executives justify need for change.
Norm test	5	New norms to judge performance
Feasibility test	7	Ways performance can be improved
Both	7	Both performance and norms
Implementation by participation	17	Stakeholder representatives determine change features.
Token	2	Solution framing with partial participation
Delegated	13	Solution specification with partial participation
Complete	2	Solution framing with full participation
Comprehensive	0	Solution specification with full participation
Implementation by persuasion	42	Experts attempt to sell a change they devise.
Consultants	13	Persuasion by consultants
Internal staff	29	Persuasion by organizational staff
Implementation by edict	23	Sponsors issue directives requiring adoption.
Change had sponsor significance	6	Sponsor has more at stake than the organization
Change had organizational significance	12	Organization has more at stake than the sponsor
Both organizational and sponsor significance	5	Both have clear-cut stakes

[a] Percentage of all cases.

---

[11] A list of each case in the data base classified by tactics, variations, and contextual factors can be obtained by writing the author.

FIGURE 3
Implementation by Intervention[a]



[a] Information flows with solid lines depict implementation-related activities and dashed lines nonimplementation-related activities. Boxes represent developmental activities, and circles represent decisions.

by applying new norms (step 2) created these rationales. Sponsor-managers often cited comparable organizations with better performance to justify new performance norms. In addition, they used descriptions of how current practices could be improved to demonstrate the feasibility of changes in step 2. These suggestions defined options that those involved in the change processes had to consider during development (steps 4 and 5). Many interventionists recognized that users were excellent sources of ideas, so during development sponsors often formed task forces made up of users to identify inefficient and ill-advised procedures. Task forces offered ideas or acted as sounding boards, providing commentaries on changes as they evolved. Sponsors, however, retained the power to veto these recommendations. Change processes ended following demonstrations that adopting changes could overcome deficiencies in performance (step 6). Performance monitoring (step 7) followed installation.

In sum, steps in planned change processes using the intervention tactic are:

1. Acquire the authority to manage a change process and appraise performance.

2. Apply new norms to identify performance inadequacies.
3. (a) Justify the new norms, or
   (b) demonstrate the feasibility of improving practices.
4, 5. Development.
6. Demonstrate improvement in performance.
7. Monitor performance.

The intervention tactic occurred in 19 percent of the cases, making it the third most frequently used approach to implementation (Table 1).

Variations. Tactical variations arose in the degree of justification that managers used. The steps in justification involved validating new norms, showing that change was possible, or both. These variations on intervention tactics will be called feasibility tests, norm tests, and dual tests.

Feasibility tests involved showing how organizational practices could be improved. For example, the manager-sponsor in one of the cases demonstrated how the organization's materials' management procedures differed from up-to-date practices of inventory control. The feasibility test was used in 7 percent of all cases and in about 35 percent of the intervention-managed cases (Table 1).

Norm tests occurred when new standards to judge performance were established. For example, a hospital CEO phased out a decentralized approach to the management of patient care. Under the decentralized system, an administrator located in each hospital ward coordinated its use of supplies and of ancillaries like x-ray; its contact with admitting, billing, other wards; and so forth. To phase out this department, the manager-sponsor contended that the administrative costs of competitors were lower because they avoided duplication of effort and diffused the accountability inherent in a decentralized system. The study found this variation in 5 percent of all cases and in about 25 percent of the intervention cases.

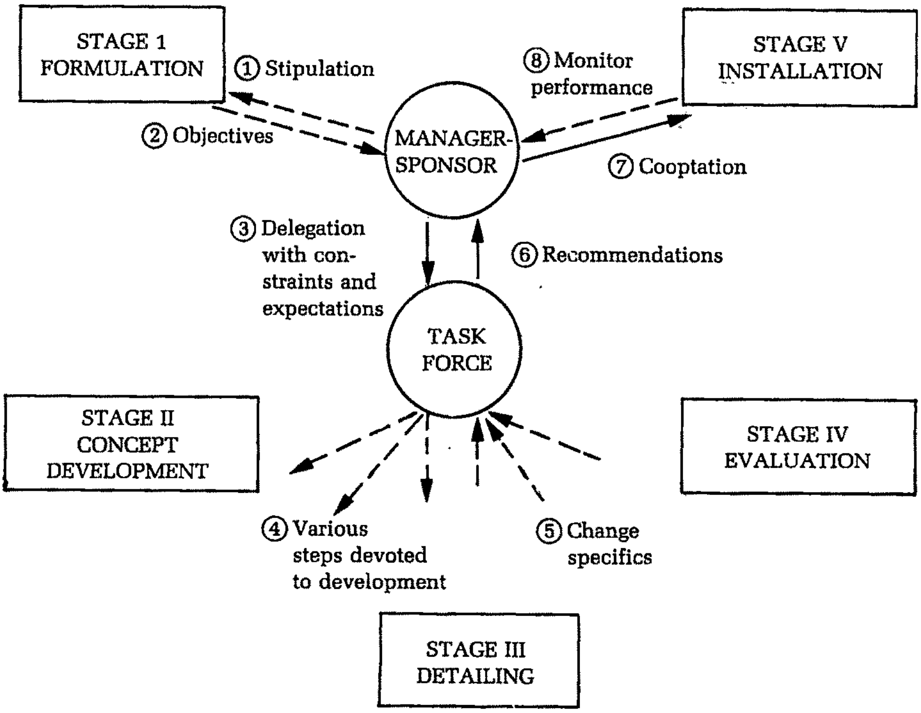
Some sponsors addressed both norms and ways practices could be improved. For example, fiscally concerned hospitals were wary of providing burn care services because third party payers, such as Blue Cross and Medicaid, did not cover all of the costs incurred. To overcome this objection one sponsor demonstrated how the cost of burn care could be covered by using a combination of sources—endowments, governmental care cost reimbursements, private insurance, and cheap resident manpower. This sponsor integrated the demonstration of financial feasibility with a description of how successful burn units were operated. The dual test variation was observed in 7 percent of all cases and about 37 percent of the cases implemented by intervention.

Implementation by Participation

The tactic. To use participation, manager-sponsors initiated change processes by stipulating needs or opportunities, and then assigning decisions for developmental activities (Figure 4, step 1). For example, a sponsor might cite the need to reduce excessive sick time or retain accreditation or opportunities to centralize activities in space vacated by another department. Articulating such stipulations (step 2) created objectives with varying degrees of explicitness, such as cutting inventory costs by 20 percent or merely reducing these costs. In step 3, sponsors selected task force members to represent certain points of view or to give information thought to be important. Sponsors delegated responsibility for guiding project development processes to task forces. Such delegation specified constraints and expectations and identified support staff. For example, in the PBX case (Figure 2) the task force had latitude to select any schedule they wished, as long as support staff determined that costs would not be increased. In steps 4 and 5, task forces arrived at consensus about specifics, such as protocols or procedures. In step 6, they made recommendations. The cooptation implicit in this delegation became a vehicle to promote compliance in step 7. The process closed with performance monitoring in step 8.

Cases were classified as participative when task forces had the authority to carry out one or more stages of a process and when leaders could not

FIGURE 4
Implementation by Participation



[a] Information flows with solid lines depict implementation-related activities and dashed lines nonimplementation-related activities. Boxes represent developmental activities and circles represent decisions.

veto the task forces' views, as in the PBX case. Sponsors made support teams available to provide information requested by task forces.

Steps in change processes using implementation by participation are:

1. Sponsors stipulate needs, opportunities, or both.
2. Sponsors set objectives.
3. Task forces formed by sponsors who
   (a) identify stakeholders,
   (b) delegate responsibility with a statement of expectations and constraints, and
   (c) assign staff support.
4, 5. Development.
6. Recommendations made to sponsor.
7. Acceptance promoted through coöptation.
8. Monitor performance.

Participation was the least frequent tactic, with 17 percent of the cases using it (Table 1).

Variations. The extent of involvement and the role of task forces defined the key variations in participation-managed implementation, as Figure 5 shows. Type 4, comprehensive participation, calls for delegation of development to fully representative task forces. This approach has the greatest cooptative potential owing to the breadth of the role and extent of participation that it allows task forces, whose memberships include all important stakeholders, such as users and power centers. Comprehensive participation is similar to Likert's System 4 (Likert, 1967). According to Likert, this approach creates considerable commitment that makes implementation likely. None of the cases used comprehensive participation, suggesting that fully cooptative approaches occur rarely, if at all, in practice.

FIGURE 5
Types of Participation

		Role of Task Force	
		Solution-framing	Solution-specifying
Extent of stakeholder involvement	Partial	Type 1: Token participation	Type 2: Delegated participation
	Full	Type 3: Complete participation	Type 4: Comprehensive participation

Type 3, complete participation, also calls for full participation but asks tasks forces to frame rather than specify developmental details. For example, users might be asked to set directions or offer ideas, with staff specialists responsible for developing changes in line with the directions set and ideas offered. The Program Planning Method also calls for this type of participation (Delbecq & Van de Ven, 1971). Complete participation seems less cooptative than the comprehensive type because it restricts the role of task forces. Changes that had long-term effects and required broad-term commitments used this type of participation. For example, in an ambulatory surgery case, all users were involved in critiquing the proposed physical structure and procedures. Complete participation appeared infrequently—in 2 percent of all cases and just 14 percent of the participation cases.

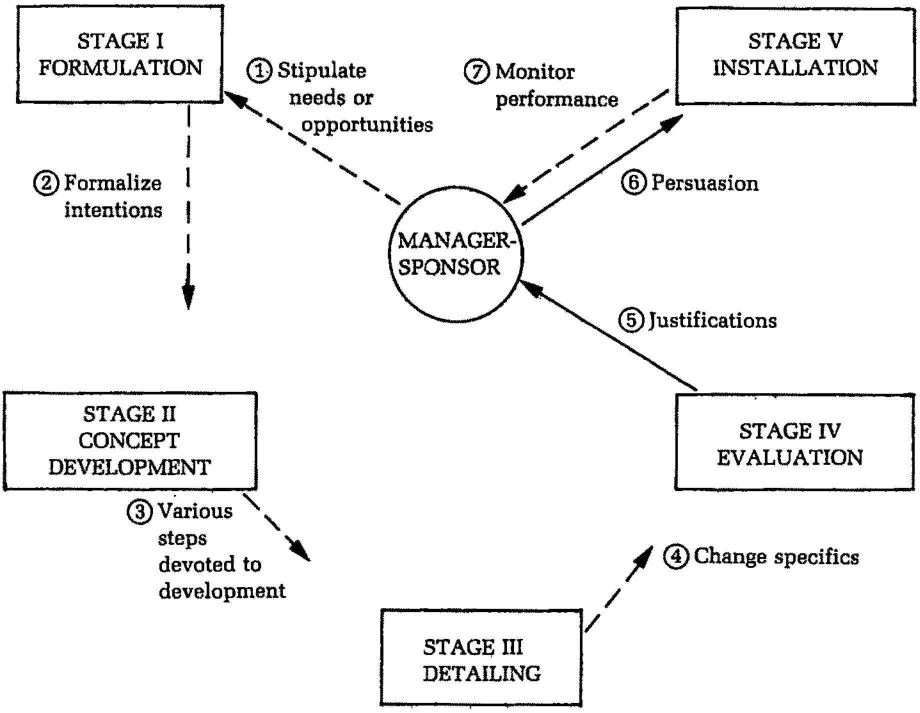
Type 2, delegated participation, uses stakeholder representatives to suggest solutions. Strategic planning groups with members drawn from organizations' boards of directors and key executives practice this type of participation (Steiner, 1979). Much of the social planning initiated in the Great Society programs of the 1960s implicitly used delegated participation to develop or modify services (Nutt, 1976). In that era, user representatives were involved in planning processes to lay out details. The PBX project used this type of participation (Figure 2). Not all the operators were involved, but participants were allowed to dictate the terms of a new schedule. In cases in this study, organizations selected benefit packages for aging programs, designed laboratories, made contingency plans for strikes, carried out renovations, and planned service procedures in this manner. In delegated participation, participants were often coopted, but their ability to persuade others to go along with changes hinged on whether the prerogatives of nonparticipants seemed threatened. For example, physicians rejected a seemingly innocuous plan for a pager system that colleagues had devised because it set a precedent of unilateral action. Thus, delegated participation may be less effective than the complete type. Delegated participation appeared in 72 percent of the participation cases and in 13 percent of all the cases (Table 1). Apparently most manager-sponsors believed that giving representatives responsibility for specifying solutions was sufficiently cooptative to gain acceptance.

The least cooptative form of participation is type 1 or token participation. Both the use of representatives and their role seemed to hamper implementation. For example, surveys may be used to frame users' problems. But problems can be misunderstood unless carefully developed sampling procedures are followed. If not, those surveyed may be unaware of stakeholder needs or misrepresent them (Nutt, 1976). Moreover, the members of representative task forces, although personally committed to changes, may fail to convince others. These two limitations suggest that token participation will be the least successful type of participation. The cases verified this speculation. Token participation was not sufficient to ensure adoption of any changes. For example, when the partners of a HMO plan disagreed about the hospital best suited to offer HMO services, Blue Cross—the most powerful member—unilaterally selected a hospital to salvage the plan. Token participation was infrequent, occurring in only 2 percent of all cases and 14 percent of the participation cases.

Implementation by Persuasion

The tactic. In implementation by persuasion, manager-sponsors made little effort to manage change processes and monitor their progress due to disinterest, lack of knowledge, or powerful or persuasive protagonists. These change processes began with either sponsors or experts stipulating needs or opportunities (step 1, Figure 6). In persuasion implementation there is little, if any, managerial review. Sponsors allow experts to control development. The path in Figure 6 moves around the periphery of the transactional model, illustrating the independence of experts, who stressed options and their justifications in steps 2, 3, and 4. Activities undertaken to pave the way for change were not observed. In step 5 experts presented products for approval. They made attempts to sell options that performed best, using

FIGURE 6
Implementation by Persuasion[a]



[a] Information flows with solid lines depict implementation-related activities and dashed lines nonimplementation-related activities. Boxes represent developmental activities, and circles represent decisions.

projected benefits to argue for adoption in step 6. Sponsors took passive roles, encouraging sales pitches from interested parties so they could weigh imperatives to act. Sponsors became active after sales pitches had been made, often demanding extensive documentation of benefits that required returning to the evaluation stage. Performance monitoring closed the process in step 7.

Steps in change processes using implementation by persuasion are:

1. Sponsors stipulate needs, opportunities, or both, or accept need or opportunity stipulations from experts.

2, 3, 4. Development.

5, 6. Interested parties use persuasion tactics to sell ideas.

7. Monitor performance.

Persuasion, the most frequently applied implementation tactic, occurred in 42 percent of the cases (Table 1).

Variations. The key variation in implementation by persuasion involved types of expert. Observation revealed two types of distinctions among experts. First was a distinction between content and process experts. Content experts were familiar with topics or systems; process experts sold procedures prescribing how to carry out processes and had no specific outcomes in mind. Content experts included consultants who sold turn-key plans to sponsors. Process experts were used when development was entrusted to organizational staffers thought to be skilled in procedures. This process-content distinction was often blurred in the cases, in part because sponsors encouraged experts to stress content (Nutt, 1984b).

The second distinction was between consultants and internal staff. Internal staff carried out development in 70 percent of the persuasion cases and 29 percent of all cases, making this variation the most widely used of all the variations in implementation tactics. When internal staff carried out processes, sponsors believed their organizations had experts skilled in either procedure or content. For example, a sponsor asked a staff member to study existing family-practice outpatient programs and recommend one to their organization. Internal staff purchased equipment, developed management procedures, planned construction and renovations, and designed accounting and other systems.

Organizations hired consultants to carry out processes in 30 percent of the persuasion cases. Consultants who were experts in specific content areas were used when manager-sponsors had little knowledge of or expertise in topics. Organizations also used consultants for projects that involved high technology, such as high-risk neonatology centers and information systems. Consultants with process expertise were used to develop construction plans and carry out specialized developmental activites, such as site selection and evaluations.

Implementation by Edict

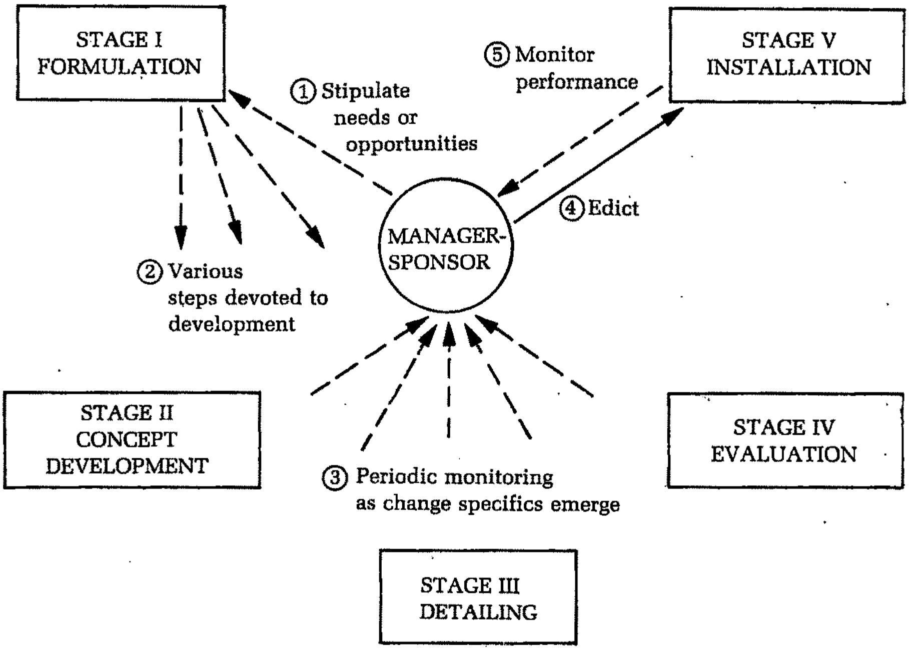
The tactic. Implementation by edict involved sponsors' using control and personal power while avoiding any form of participation. This tactic had three key features. First, sponsors' control of change processes was intermittent with no common theme. There was no recognizable pattern in sponsors' monitoring and no formal delegations of responsibility. Experts and users had little power. Second, sponsors did not discuss changes with users or attempt to rationalize the need for changes; they expected user compliance. They demonstrated neither imperatives for changes nor their feasibility. The justification steps found in the intervention tactic were not observed. Third, sponsors issued adoption directives by managerial fiat even when lengthy developmental processes occurred. Sponsors merely announced changes and prescribed expected behavior using memoranda, formal presentations, or on-the-job instructions that dictated the expected behavior of users. Sponsors applying power in this way used implementation by edict.

The steps associated with edict implementations in change processes shown in Figure 7 are:

1. Sponsors stipulate needs or opportunities.
2, 3. Development.
4. Sponsors issue directives.
5. Monitor performance.

Edicts were used in 23 percent of the cases (Table 1).

FIGURE 7
Implementation by Edict[a]



[a] Information flows with solid lines depict implementation-related activities, and dashed lines nonimplementation-related activities. Boxes represent development activities, and circles represent decisions.

Variations. The study identified three notable variations in the use of power tactics. These variations were based on whether changes had individual or organizational significance, or both. Sponsors had personal vested interests in several of the edict-implemented changes. Some changes were identified with sponsors to the extent that failures to adopt them would have hindered these managers' upward mobility. For example, a hospital CEO, hoping to alleviate disturbing malpractice incidents, announced a patient relations department by having the head of this new department make a presentation to key executives. Other sponsors saw themselves as acting in their organizations' best interests by issuing edicts because they thought the changes had special significance. For instance, a sponsor used a memorandum to announce a new pricing policy for an emergency room, hoping to stem recent losses of business to freestanding urgent care centers. Still others believed that their interests and their organizations' interests were synonymous. In the case of a helicopter transport system, for example, both the plans of a key competitor and the CEO's pet idea provided imperatives to act.

In about 25 percent of the cases of implementation through edicts and in 6 percent of all cases sponsors had more at stake than their organizations. Changes had clear-cut organizational significance with modest or minimal implications for sponsors in half of the edict cases and 12 percent of all cases. Organizations and sponsors shared equal stakes in changes in about 25 percent of the edict-implemented cases and in 5 percent of all cases. Sponsors were found to have important stakes in half of the cases of implementation by edict.

Combined Tactics as Hybrids

In hybrid cases, sponsors switched implementation tactics. Two or more of the basic tactics appeared in these cases. For example, some sponsors took responsibility away from experts, creating persuasion-intervention hybrid tactics. Other combinations included participation-intervention and persuasion-participation. The tactics used, their order, and their relative emphasis all seem important. However, there were too few hybrid cases to show if the use of combinations meant managers were groping for correct tactics, with the last tactics used being their choice, or if the combinations themselves gave improved leverage. Hybrids may signal managerial errors or the emergence of different implementation needs at different stages of development. Only 7 of 91 cases, or 7 percent of the cases in the data base, were classified as hybrids, suggesting that the type is rare.

SUCCESS OF TACTICS AND CONDITIONS OF USE

Adoption Rates

The study determined tactic success by treating the tactic options in Table 1 as independent variables and adoption as the dependent variable.

ANOVA results indicated that both tactics and tactical variations were correlated with adoption ($p < .005$ and $p < .05$, respectively).

The intervention tactic had a 100 percent success rate, followed by participation with 84 percent, persuasion with 73 percent, and edicts with 43 percent. A Duncan multiple range test found that the success rates for the intervention and edict tactics were significantly different. The persuasion and participation tactics were similar but distinct from the intervention and edict tactics ($p < .05$). This suggests that intervention has the best prospects of success, edict the worst, and persuasion and participation similar, intermediate prospects. The success rates for the tactical variations followed this ordering: Tactical variations for the intervention tactic had success rates that were better than or equal to those for variations on the participation tactics; variations on participation tactics were better than or equal to the persuasion variations; and the persuasion variations were better than the variations on implementation through edicts.

The most successful tactical variations were (1) the intervention variations, including use of norm, feasibility, or both norm and feasibility tests, and (2) the token and complete participation variations (Table 2). If adoption is considered the measure of success, the intervention tactical variations seem clearly superior to all others. The success of token and complete participation is difficult to interpret because there were only two cases for each type. Delegated participation, or solution-seeking by representatives, had success rates similar to those of the persuasion variations, according to the Duncan multiple range test. This decline in rate of success suggests that participation may lose its cooptative power unless task forces involve most, if not all, of the people whom changes affect. When representatives were used, participation became much like persuasion, both in terms of approach—representatives selling changes—and likelihood of success. All of the edict variations fared poorly, with success ranging from 30 to 60 percent.

Contextual Factors and Conditions of Use

Table 2 shows the values obtained for the contextual factors sorted by tactic and variation. The analysis of these contextual factors was carried out to determine relationships between conditions surrounding changes and use of implementation tactics. Specifying these conditions helped to qualify the success rates observed for the tactics and their variations. In this analysis, perceived importance and time pressure associated with changes did not lead to the use of particular tactics or tactical variations, but budgets for change processes and staff support did seem to influence managers' selection of tactics.

When variations of intervention or persuasion tactics were used, process budgets were above average ($p < .02$). Delegated participation, or development of solutions using representatives, had process budget demands below those of other changes. Both token participation, or framing of solutions using representatives, and implementation by edict for changes with high significance to sponsors also had low budget demands. Perhaps sponsors

TABLE 2
Performance and Characteristics of the Implementation Tactics

Tactics and Variations	Number of Cases	Success Rate	Contextual Factors			Types			
			Perceived[a] Importance	Staff[b] Support	Time[c] Pressure	Process[d] Budget	Program	Construction	Equipment
Intervention									
Norm test	4	100%	3.50	4.00	3.50	4.00	50%	50%	0
Feasibility test	6	100%	2.84	3.50	3.00	3.67	66%	17%	17%
Both tests	6	100%	3.17	2.80	2.67	3.84	84%	16%	0
Overall success	16	100%	3.13	3.40	2.88	3.69	69%	25%	7
Participation									
Token	2	100%	3.00	2.50	3.00	2.00	100%	0	0
Delegated	10	80%	2.78	3.45	1.89	2.78	60%	30%	10%
Complete	2	100%	3.00	3.50	2.00	5.00	0	50%	50%
Comprehensive	None								
Overall success	14	84%	2.85	3.31	2.08	3.00	58%	29%	15%
Persuasion									
By staff	24	74%	2.67	3.59	2.34	3.30	55%	21%	25%
By consultants	11	64%	3.40	3.60	2.40	4.00	64%	28%	9%
Overall success	35	73%	2.89	3.59	2.36	3.47	58%	23%	20%
Edict									
Sponsor significance	5	60%	2.20	3.60	2.20	2.20	100%	0	0
Organizational significance	10	30%	2.75	2.75	2.00	3.13	60%	40%	0
Both	4	50%	2.75	3.25	3.25	3.75	25%	0	75%
Overall success	19	43%	2.59	3.12	2.36	3.00	64%	21%	16%
Totals	84	71%*	2.87	3.37*	2.42	3.34**	61%	24%	16%

[a] Scale points are: 4 = critical, 3 = very important, 2 = important, 1 = routine.
[b] Scale points are: 4 = good, 3 = satisfactory, 2 = marginal, 1 = poor.
[c] Scale points are: 4 = very high, 3 = high, 2 = moderate, 1 = low, 0 = absent.
[d] Scale points are: 4 = far above, 3 = similar, 2 = below, 1 = far below.
* $p < .05$
** $p < .02$

kept the budgets of changes significant to them and implemented by edicts low to hide their interests. Large process budgets might draw unwanted attention. Cases of token participation may be pro forma attempts at involvement designed to minimize costs, which may explain their relatively low process budgets. More cases are needed to explain the high success rates and low costs associated with token participation.

Cases that used complete participation, or fully representative solution framing, had budgets that were far above those of typical changes. Using task forces that were fully representative to define problems or comment on plans resulted in large budgets, but was effective, as the 100 percent adoption rate for this variation indicates. Costs of involving large numbers of people may account for these comparatively high process budgets. Delegated participation, with solutions framed by representatives, was less costly but also less effective because task force members had to resort to persuasion to sell their ideas. Persuasion by several task force members succeeded only a bit more often than persuasion by single experts, as the data in Table 2 verifies.

The quality of staff support was related to the tactical variations ($p < .05$). Staff support perceived to be satisfactory to good characterized all but organizationally significant edicts and token participation (framing with representatives). Respondents designated support as marginal for organizationally significant edicts and as just below satisfactory for token participation. This suggests that sponsors using edicts may underestimate the need for good staff support when changes have organizational significance but do not when changes have personal significance. The personal involvement of sponsors heightens attention and entices them to recruit capable people to help carry out changes. When such heightened interest was lacking, sponsors used staff thought to be less capable.

Sponsors seemed to underestimate needs for staff support when they used delegated participation, or representatives devising solutions. They may have believed that task forces were self-sufficient and failed to see the need for staff to support development of solutions. The demands of solution development were high here compared to the demands created by framing solutions in token and complete participation.

The study also classified changes as equipment, construction, and program types, as discussed earlier in this paper. Equipment changes often captured the attention of key administrators because they typically cost $500,000 or more. Development focused on determining the value of equipment innovations or updating. Construction projects involved changes in the physical features of key facilities. Development stressed physical features and rationales that supported needs to alter, add to, or build new facilities. Large budgets of 10 million and more and significance throughout organizations were typical. The category program change described installing new services or internal operations that involved minor equipment and construction expenditures, if any, of less than a quarter of the budget. Most program changes had only incidental equipment or construction costs. The emphasis was to modify or devise new institutional practices that developed software, procedures, protocols, and the like for services like detoxification units, or for internal operations, like security programs. These changes stressed procedural development and allocated the bulk of their budgets for this purpose.

Program changes were the most common, representing 60 percent of the cases. About 25 percent were construction changes and 16 percent were equipment changes. Unlike Damanpour and Evan (1984), this study found no relationship between implementation tactics and type of change. Managers do not appear to favor particular tactics to implement program, equipment, or construction changes.

CONCLUSIONS

This research led to several insights about the practice of implementation that offer guidelines for managers and suggest topics meriting further research. Key conclusions concern implementation practices and the effectiveness of implementation tactics under particular conditions of use.

The Nature of Managerial Implementation

The intervention tactic. The intervention tactic is similar to the change agent prescriptions of Lewin (1947) and Schein (1964), with some important distinctions. In practice, change agents were manager-sponsors who took control of planned change processes. Consultants were never observed. Steps were similar to the Lewin-Schein implementation procedure of unfreezing, change, and refreezing (Dalton, 1970). However, managers in this study were found to be far more aggressive than organizational development (OD) specialists acknowledge as either necessary or desirable. Behaving in line with the prescriptions of Kotter and Schlesinger (1979), successful implementers carefully monitored entire change processes, regulating and controlling social and political issues as they arose. Managers using the intervention tactic were quite good at creating new norms in systems they sought to change. They offered new definitions of acceptable performance, justified these new norms, and showed how practices could be improved. Through these steps, sponsors were able to stipulate needs, thereby eliminating ambiguity, deal with scale, self-educate, collect essential data to deflect threats, manage resentment, build confidence, and reinforce what was wanted. These are all key forces identified by Zand and Sorensen (1975) in their study of implementation in the Lewin-Schein process frame. Only one of the managers had any exposure to the OD literature, suggesting that these practices evolved from practice-based heuristics, as suggested by Schön (1983).

The participation tactics. The origins of participation as an implementation tactic can be traced to the classic study by Coch and French (1948), who found that people reacted more favorably and became more committed when they participated in change-making processes than when they did not. This tactic mandates forming task forces that represent key stakeholders. Manager-sponsors delegate their development prerogatives to the task forces. The Delbecq, Van de Ven, and Gustafson (1975) approach to planned change is typical of many contemporary approaches that make these prescriptions.

Managers who used participation tactics were aware of the participation literature and often said that they used participation because of this knowledge. However, the study observed no instance of full or comprehensive participation and thus cannot specify the details of such a tactic. Typically, managers scaled down the scope of participation. Representatives acted as agents for stakeholders in more than 85 percent of the participation-implemented projects. In those few cases where full involvement occurred, participants had narrow roles. They commented on issues that helped frame problems by interpreting their own needs or critiquing proposed solutions.

The benefits of participation seem contingent on what task forces are asked to do and the scope of the involvement of key people. Representatives can help, but the power of cooptation drops dramatically for those not directly involved. The effects of cooptation decline both when representatives of a large group of stakeholders are involved and when the role of a task force is limited. All tactics but comprehensive participation have these limitations. However, comprehensive participation, or full representation of stakeholders in devising solutions, was not observed and may seldom, if ever, arise in practice owing to its cost. Even the complete participation type, full participation of stakeholders for a part of the process, can be difficult to carry out because affected parties are often numerous and diffuse. The parties involved in a prison siting are an example. The leverage from delegated participation stemmed from representatives, but the number of individuals affected by the change that could be involved limited the power of cooptation. Participation should be more carefully defined in future research to account for participants' roles and the extent of their involvement.

The persuasion tactic. Operations researchers and others who call for experts to guide development activities in planned change processes advocate a sales approach to implementation. For example, Churchman (1979b) argued that implementation depends on anticipating and countering moral, aesthetic, religious, and political objections to proposed changes. According to these views, implementation should hinge on experts who determine what should be done and use rational arguments to convince managers to go along. Organizations like the National Science Foundation, the Ford Foundation, and the National Institute of Health apply this type of implementation in their peer review procedures. Experts select the grants and contracts they will support—their products. The controversial Operation Research Society of America guidelines, published in 1971, also called for experts to vigorously sell ideas they find desirable. As a result, many of the planned change processes devised by management scientists and operations researchers have experts recommending and managers reacting. Implementation is seen as an educational process (Churchman, 1975). When changes are not adopted, failure is often attributed to managers who did not understand developmental processes or their products (Churchman & Schainblatt, 1965; Duncan, 1974).

Sponsors extensively delegated development to experts. More than 40 percent of the cases used expert-managed development processes. Managers gave experts broad responsibility, but withheld approval until the experts could demonstrate value. This forced experts to devise well-constructed rationales. The frequency with which persuasion tactics were used to justify decisions to adopt changes suggests that many sponsors see development as an activity that internal staff or consultants should carry out. Forcing experts to gather arguments for adopting changes at the expense of gaining the acceptance of stakeholders may account for many of the failures of implementation by persuasion.

Future research should consider how consultants influence persuasion-based processes. Sponsors can use consultants or internal staff, but seem to prefer internal staff. Consultants play vital roles in many organizations, but their activities and the benefits they create are shrouded in secrecy. More study of consultants and their role in implementation seems essential.

The edict tactic. Issuing edicts requires sponsors to use one or more of the reward, legitimate, expert, informational, and referent power types identified by French and Raven (1959). The authority to offer rewards or to remove irritants is called reward power. To apply this type of power, managers must have control over financial and nonfinancial incentives that can be used to entice others to adopt plans. Legitimate power is based on sponsors' rights to take action in organizations. Expert power flows from a track record of successes and a reputation for candid and honest dealings. Expert power requires both credibility and trust, but informational power, in which managers offer their insights or cite anecdotes, requires only that they have the trust of other organizational members. Referent power stems from the attractions people have for others that lead them to label these individuals as charismatic. Sponsors drew on one or more of these power bases when carrying out edicts (Nutt, 1983). For example, legitimate power was applied when sponsors used legitimate prerogatives to make changes. When carrying out edicts, sponsors trade social credit for action and risk failure if their power proves to be insufficient. However, the frequent use of power can strain organizations and gradually drain manager-sponsors' stores of social credit.

Contingency use of tactics. Research is needed to determine if contingency use of implementation tactics is desirable. All forms of implementation were found to have some degree of success. Because some ways to manage planned change processes are more costly and less timely than others, future research could develop and test conditions in which each implementation tactic can be effective, creating a contingency model.

Tactic Effectiveness and Conditions of Use

The study identified a variety of interventionist tactics such as challenging existing norms, offering viable ideas, or both, and found all to produce 100 percent success rates. The most common participation tactic—representatives providing solutions—implicitly used persuasion-like tactics and had success rates of about 75 percent, similar to those for persuasion tactics.

Edicts fared poorly, failing more often than they succeeded, whether changes had personal significance for sponsors, organizational significance, or both.

The only participation variation that employed full representation of affected parties occurred infrequently, perhaps because complete participation demanded process budgets far above what other changes required. Generating solutions with delegated participation required modest process budgets, but was less effective than complete participation, seemingly because representatives had to use persuasion to sell changes they devised. Managers underestimated needs for staff support in delegated participation. Apparently, sponsors allocate too little staff support when they enlarge the role of task forces to include solution development. Token participation, framing with representatives, had marginal staff support and small process budgets, suggesting pro forma attempts at cooptation and efforts to minimize implementation costs.

Edicts motivated by sponsors' interests had small budgets, apparently to avoid drawing undue attention to changes and sponsors' vested interests. When an edict was used for changes with organizational significance, managers provided only marginal staff support, but respondents termed staff support satisfactory for changes of some significance to sponsors. Sponsors appear to reserve the best staff for their pet change attempts.

Success was guaranteed when sponsor-managers were actively involved in applying interventionist tactics. The 100 percent success rates for intervention and its modest resource demands make this implementation tactic highly desirable and worthy of active promotion among managers. Persuasion was successful in three-quarters of the cases in which managers used it. This tactic demanded strong staff support and seemed to require high process budgets compared to intervention. Surprisingly, implementation by persuasion was not limited to changes that sponsors considered of little importance and under low time pressure. Managers did not deem only throw-away changes appropriate for experts to manage, as some theorists have implied (e.g., March, 1981). Participation was found to be no more effective than persuasion, and the two required comparable staff and process budget support.

All of the tactics were used under varying levels of time pressure and perceived importance and for each of the change types. This suggests that time pressure, importance, and change type do not influence managers' selections of implementation tactics.

Future researchers should be wary of pragmatic measures, such as the one used in this research, rate of adoption. The PBX case demonstrates the reason for this caveat. What was the origin of the grievances? At the conclusion of the change process no one had any idea. The sponsor saw the symptoms disappear and so judged the change a success. But if the origin of the grievances were an autocratic PBX supervisor, the grievances might recur. What appeared to be a pragmatic change may have merely papered over issues that could emerge again in problems seemingly unrelated to the earlier ones. Implementation has pragmatic importance, but the key step in planned change processes may be formulation. Future research should devise ways to measure the likelihood and consequences of incorrect stipulation of needs or opportunities by managers.

REFERENCES

Ansoff, H. I. 1984. ***Implanting strategic management.*** Englewood Cliffs, N. J.: Prentice Hall International.

Argyris, C. 1970. Resistance to rationale management systems. ***Innovation***, 10: 28–34.

Bardach, E. 1977. ***The implementation game.*** Cambridge, Mass.: MIT Press.

Bartlett, F. C. 1954. ***Remembering: A study in experimental and social psychology.*** Cambridge, Mass.: Harvard University Press.

Bennis, W. 1973. ***Beyond bureaucracy.*** New York: McGraw-Hill Book Co.

Beyer, J. M., & Trice, H. M. 1978. ***Implementing change: Alcoholism policies in work organizations.*** New York: The Free Press.

Bourgeois, L. J., & Brodwin, D. R. 1984. Strategic implementation. ***Strategic Management Journal***, 5: 241–264.

Bower, J. L. 1970. ***Managing the resources allocation process: A study of corporate planning and investment.*** Homewood, Ill.: Richard D. Irwin.

Campbell, D. T., & Fiske, D. W. 1959. Convergent and discriminant validation by the multitrait-multimethod matrix. ***Psychological Bulletin***, 56: 81–105.

Churchman, C. W. 1975. Theories of implementation. In R. L. Schultz & D. P. Slevin (Eds.), ***Implementing operations research/management science***: 23–30. New York: Elsevier.

Churchman, C. W. 1979a. ***The systems approach.*** New York: Dell.

Churchman, C. W. 1979b. ***The systems approach and its enemies.*** New York: Basic Books.

Churchman, C. W., & Schainblatt, A. H. 1965. The researcher and the manager: A dialectic of implementation. ***Management Science***, 11 (4): B69–B87.

Coch, L., & French, J. R. P., Jr. 1948. Overcoming resistance to change. ***Human Relations***, 1: 512–532.

Daft, R., & Becker, S. W. 1978. ***Innovation in organizations.*** New York: Elsevier.

Dalton, G. W. 1970. Influence and organizational change. In G. W. Dalton, P. Lawrence, & L. Greiner (Eds.), ***Organizational change and development***: 230–258. Homewood, Ill.: Richard D. Irwin.

Damanpour, F., & Evan, W. M. 1984. Organizational innovation and preference: The problem and organizational lag. ***Administrative Science Quarterly***, 29: 392–409.

Delbecq, A., & Van de Ven, A. 1971. A group process model for problem identification and program planning. ***Journal of Applied Behavioral Science***, 7: 466–492.

Delbecq, A., Van de Ven, A., & Gustafson, D. H. 1975. ***Group techniques for program planning.*** Glenview, Ill.: Scott-Foresman & Co.

Duncan, W. J. 1974. The researcher and the manager: A comparative view of the need for multiple understanding. ***Management Science***, 20: 1157–1163.

Dunn, W. N. 1981. ***Public policy analysis***: Englewood Cliffs, N.J.: Prentice-Hall.

French, J., & Raven, B. 1959. The bases of social power. In D. Cartwright (Ed.), ***Studies in social power***: 150–167. Ann Arbor, Mich.: Institute for Social Research.

Ginsberg, M. J. 1975. ***A process approach to management science implementation.*** Unpublished doctoral dissertation, Massachusetts Institute of Technology, Cambridge.

Glass, G. V., McGaw, B., & Smith, M. 1982. ***Meta-analysis in social research***. Beverly Hills, Calif.: Sage.

Greiner, L. E. 1970. Patterns of organizational change. In G. W. Dalton, P. R. Lawrence, & L.E. Greiner (Eds.), ***Organizational change and developments***: 213–229. Homewood, Ill.: Richard D. Irwin.

Gustafson, D., Shukla, R., Delbecq, A., & Wallster, G. 1973. A comparative study in subjective likelihood estimates made by individuals, interacting groups, delphi groups, and nominal groups. ***Organizational Behavior and Human Performance***, 9: 280–291.

Hage, J., & Aiken, M. 1970. ***Social change in complex organizations***. New York: Random House.

Keen, P., & Scott-Morton, M. 1978. ***Decision support systems: An organizational perspective***. Reading, Mass.: Addison-Wesley.

Kotter, J. P., & Schlesinger, C. A. 1979. Choosing strategies for change. ***Harvard Business Review***, 57 (2): 16–114.

Lewin, K. 1947. Group decisions and social change. In J. E. Maccoby, T. W. Newcomb, & E. Harley (Eds.), ***Readings in social psychology***: 459–493. New York: Holt, Rinehart & Winston.

Likert, R. 1967. ***The human organization***. New York: McGraw-Hill Book Co.

Lucas, H. 1975. ***Why information systems fail***. New York: Columbia University Press.

March, J. G. 1981. Footnotes to organizational change. ***Administrative Science Quarterly***, 26: 563–577.

McKelvey, B. 1978. Organizational systematics: Taxonomic lessons from biology. ***Management Science***, 24: 1428–1440.

Mintzberg, H., Raisinghani, D., & Theoret, A. 1976. The structure of unstructured decision processes. ***Administrative Science Quarterly***, 21: 246–275.

Mintzberg, H., & Waters, J. A. 1982. Tracking strategy in an enterpreneural firm. ***Academy of Management Journal***, 25: 465–499.

Nutt, P. C. 1976. The merits of using consumers as members of planning groups. ***Academy of Management Journal***, 19: 378–394.

Nutt, P. C. 1979. Calling out and calling off the dogs: Managerial diagnosis in organizations. ***Academy of Management Review***, 4: 203–214.

Nutt, P. C. 1983. Implementation approaches for planning. ***Academy of Management Review***, 8: 600–611.

Nutt, P. C. 1984a. ***Planning methods***. New York: John Wiley & Sons.

Nutt, P. C. 1984b Types of organizational decision processes. ***Administrative Science Quarterly***, 29: 414–450.

Pressman, J., & Wildavsky, A. 1973. ***Implementation***. Berkeley: University of California Press.

Schein, E. H. 1961. ***Coercive persuasion***. New York: Norton.

Schein, E. H. 1964. The mechanisms of change. In W. Bennis, E. Schein, G. Stelle, & A. Berlow (Eds.), ***Interpersonal dynamics***: 362–378. Homewood, Ill.: Dorsey Press.

Schendel, D., & Hofer, C. (Eds.) 1979. ***Strategic management***. Boston: Little, Brown & Co.

Schon, D. A. 1983. ***The reflective practitioner: How professionals think in action***. New York: Basic Books.

Schultz, R. L., & Slevin, P. O. 1975. ***Implementing operations research/management science***. New York: Elsevier.

Soelberg, P. O. 1967. Unprogrammed decision making. ***Industrial Management Review***, 8 (2): 19–29.

Steiner, G. 1979. ***Top management planning***. New York: Macmillan Book Publishing Co.

Waters, J. A., Salipante, P. F., Jr., & Notz, W. W. 1978. The experimenting organization: Using the results of behavioral science research. ***Academy of Management Review***, 3: 483–492.

Witte, E. 1972. Field research on complex decision making processes—the phase theory. ***International Studies of Management and Organization***, (Fall) 156–182.

Yin, R. K. 1981. The case study crisis: Some answers. ***Administrative Science Quarterly***, 26: 58–65.

Zaltman, G., Duncan, R., & Holbek, J. 1973. ***Innovations in organizations***. New York: Wiley-Interscience.

Zand, D. E., & Sorensen, R. E. 1975. Theory of change and the effective use of management science. ***Administrative Science Quarterly***, 20: 532–545.

Paul C. Nutt received his Ph.D. degree at the University of Wisconsin at Madison and is a professor in the departments of Hospital and Health Services Administration; Management and Human Resources; and Management Science at the Ohio State University, Columbus. His current research interests include strategic management, planning, and decision making.

© *Academy of Management Journal*
1986, Vol. 29, No. 2, 262–279.

SOCIALIZATION TACTICS, SELF-EFFICACY, AND NEWCOMERS' ADJUSTMENTS TO ORGANIZATIONS

GARETH R. JONES
Texas A & M University

This study investigated the relationship between the socialization tactics employed by organizations and a series of role and personal outcomes. It also examined the effects of self-efficacy on role orientation. Consistent with predictions, results suggest that different patterns of socialization lead to different forms of newcomer adjustment to organizations. Specifically, institutionalized tactics lead to custodial role orientations, and individualized tactics to innovative role orientations. The results also suggest that self-efficacy moderates this learning process—specifically that socialization tactics produce a stronger custodial role orientation when newcomers possess low levels of self-efficacy.

Although there is evidence to suggest that socialization experiences affect personal and role outcomes, few empirical studies of socialization processes have explicitly addressed the question of how specific socialization tactics affect newcomers' adjustments to organizations. Rather, the focus has been on how newcomers' experiences of organizational reality at different stages in socialization processes result in such outcomes as satisfaction and feelings of personal worth (Feldman, 1976), commitment (Buchanan, 1974), or longevity in organizations (Katz, 1978). Only untested theories have explicitly addressed the way in which socialization tactics affect newcomers' responses. A second and related issue is the possibility that individual differences may also affect newcomers' adjustments to organizations and may furthermore moderate the effects of socialization tactics on individuals' orientations to roles (Jones, 1983a). From such an interactionist perspective, it is necessary to investigate how both individual and organizational factors combine to influence (1) newcomers' subsequent adjustments to organizations and (2) their role behaviors.

The purpose of this study was to address these issues. It first examined how the information provided by organizations through their socialization practices may influence the way newcomers adjust to the organizations. Second, it examined the proposition that newcomers' levels of self-efficacy

---

The author wishes to express his appreciation to Lyle Schoenfeldt, Tom Bateman, Gerry Ferris, and Cynthia Fisher for their helpful comments on earlier drafts of this manuscript. I also wish to thank Jane Nathan and the College of Commerce and Business Administration at the University of Illinois for providing resources to conduct the study. An earlier version of the manuscript was presented at the Western Academy of Management Meetings, March 1984.

may moderate the effects of these socialization practices on their subsequent role orientations. Role orientation is defined here as the manner in which individuals perform their roles and adjust to task requirements.

TACTICS AND OUTCOMES

Newcomers entering organizations may experience a reality shock (Hughes, 1958) or surprise (Louis, 1980) when their assumptions about how people interpret and respond to actions or events do not conform with those that prevail in their new contexts. As a result, newcomers may be forced to reevaluate their assumptions about these organizations and to seek information about why people behave as they do, if they are to reduce the uncertainty or anxiety surrounding the entry process (Jones, 1983b; Van Maanen, 1977). One of the main ways in which organizations can influence this learning process is by using specific tactics of socialization.

Van Maanen and Schein (1979) offered a theoretical explanation of how methods of socialization influence one particular outcome of socialization, role orientation. According to their theory, newcomers respond to their roles differently because the socialization tactics used by organizations shape the information newcomers receive. By giving or withholding information or by providing information in particular ways, organizational incumbents can encourage newcomers to interpret and respond to situations in a predictable manner. Figure 1 presents the six socialization tactics that Van Maanen and Schein (1979: 236–252) hypothesized to be significant in influencing

FIGURE 1
A Classification of Socialization Tactics[a]

Tactics concerned mainly with:	INSTITUTIONALIZED	INDIVIDUALIZED
CONTEXT	Collective Formal	Individual Informal
CONTENT	Sequential Fixed[b]	Random Variable
SOCIAL ASPECTS	Serial Investiture[b]	Disjunctive Divestiture

[a] Based on Van Maanen and Schein (1979: 232).
[b] Indicates reverse of effects hypothesized by Van Maanen and Schein (1979).

newcomers' responses. However, the present study hypothesized that the effects of two of these tactics on newcomer response are the opposite of those proposed by Van Maanen and Schein. Also, the overall classificatory scheme used in this study is different, as is discussed in the following pages.

The first two types of tactics vary in terms of the contexts in which organizations provide information to newcomers. First, collective versus individual socialization tactics are hypothesized to produce divergent role orientations. With collective tactics, new recruits go through common learning experiences designed to produce standardized responses to situations. In group settings interactions among newcomers reinforce the definition of the situation offered by socialization agents. By contrast, individual socialization practices give each newcomer a unique set of learning experiences that allows heterogeneity in their responses. As a result, collective tactics will produce custodial role orientations, in which newcomers accept the status quo and passively accept the substantive requirements of tasks or roles (Van Maanen & Schein, 1979). Individual tactics, on the other hand, provide newcomers with opportunities to develop differentiated responses and to adopt innovative orientations toward roles. With such orientations, newcomers may decide to change the methods for performing their roles or even their missions.

The second categorization of socialization tactics, formal versus informal, is closely related to the first in that it also deals with the context in which information is presented to newcomers. When organizations employ formal tactics, they segregate newcomers from other organizational members while they learn the responsibilities of their roles. With informal tactics, newcomers become part of work groups and learning takes place on the job. As a result, formal tactics, especially when coupled with collective practices, increase the propensity of newcomers to accept definitions of situations offered by significant others in their organizations. In turn, this increases the degree to which newcomers will share common norms, values, and attitudes, and develop custodial orientations. By contrast, informal and individual tactics provide newcomers with great latitude to make differentiated responses, and innovative responses are a likely result.

The next two categorizations of tactics, sequential versus random and fixed versus variable, deal with the content of the information given to newcomers via socialization. Sequential tactics provide recruits with explicit information concerning the sequences of activities or experiences they will go through in their organizations. Fixed tactics provide them with precise knowledge of the timetables associated with completing each stage in these processes. By contrast, variable socialization tactics provide no information about when newcomers may reach a certain stage in a learning process and, when a process is random, they do not know the sequence of its stages. Thus, random and variable practices will increase newcomers' levels of uncertainty since they are missing information concerning their organizational futures.

While Van Maanen and Schein (1979) argued that sequential tactics will lead to role conformity, they also argued that variable tactics will lead to custodial responses because variability causes anxiety that motivates people strongly towards conformity. However, it seems more likely that when the future is uncertain, variable tactics will encourage innovative responses, especially in organizations in which newcomers' abilities to deal competently with uncertain situations govern upward mobility. For similar reasons, fixed socialization tactics may not lead to innovative responses, but to custodial ones, because individuals will have no desire to rock the boat if they can clearly see the pathways to their future statuses from the beginning. Thus, the effects of fixed or variable tactics on newcomers' orientations must be determined empirically.

Van Maanen and Schein's (1979) last two categorizations of socialization tactics are also closely related to each other; they are serial versus disjunctive and investiture versus divestiture. In serial processes, experienced organizational members act as role models for new recruits, but in disjunctive processes newcomers must develop their own definitions of situations because no other or prior role incumbents are available. For this reason, disjunctive processes are likely to result in active, innovative role orientations. The final tactical category, investiture versus divestiture, concerns the degree to which newcomers receive positive or negative social support after entry from experienced organizational members. These two categorizations therefore reflect social or interpersonal aspects of the socialization process. In the social, interpersonal sphere, interpretations offered by other organizational members may more strongly influence newcomers' perceptions of contexts than do objective characteristics of contexts. Fixed or sequential socialization tactics would furnish such objective information. Previous research (Van Maanen & Schein, 1979) suggested that divestiture also results in custodial responses because organizations explicitly intend such tactics to mold newcomers into forms that the organizations wish. By contrast, investiture, which reinforces peoples' beliefs in their own competency, results in innovative role orientations.

Once again, the logic behind this argument is open to question. By disconfirming newcomers' expectations about themselves through negative social experiences, divestiture may not cause newcomers to adopt passive or custodial role orientations. Instead, it may encourage them to excel in their roles. In other words, divestiture may lead individuals to question definitions of situations offered by others and stimulate innovative orientations to roles. Conversely, confirming newcomers' definitions of their own competency at an early stage in their entry processes may paradoxically cause them to believe in self-fulfilling prophecies of their own worth that will lead to custodial orientations. Thus it is not clear a priori which socialization tactics will result in active and innovative or passive and custodial role orientations. Furthermore, it is unclear whether different sets of practices, as opposed to different ends of the six continua, will lead to opposite role orientations. Van Maanen and Schein (1979) suggested that different combinations of tactics will result in different role orientations. However, given the modifications suggested above and the lack of previous empirical work, the guiding hypothesis for this study is that different ends of the six continua will produce opposite role orientations. Specifically, it was expected that socialization practices that are individual, informal, variable, random, and disjunctive and that involve divestiture will produce innovative role orientations. These practices will be collectively referred to as *individualized* socialization tactics. Conversely, it was expected that the opposite ends of these continua, collectively referred to as *institutionalized* socialization tactics, lead to passive or custodial role orientations.

> *Hypothesis 1: Institutionalized socialization tactics will produce custodial role orientations and individualized tactics will produce innovative role orientations.*

Role orientation is only one outcome that the process of socialization may influence. Other outcomes of importance concern levels of role ambiguity and role conflict, job satisfaction, commitment, and intention to quit. Taken collectively, these outcomes generally describe newcomers' personal adjustments to their organizations. To the degree that socialization tactics provide information that reduces the uncertainty or anxiety surrounding an entry process, a predictable pattern of relationships between these variables should emerge. Specifically, since institutionalized socialization tactics are likely to present newcomers with less problems in searching for situational consistency and mediating personal adjustment than individualized tactics, it is hypothesized that:

> *Hypothesis 2: Institutionalized socialization tactics will be negatively related to role conflict, role ambiguity, and intention to quit, and positively related to job satisfaction and commitment.*

Although Hypothesis 2 implies that, in general, institutionalized tactics are more likely to influence personal adjustments to organizations than individualized tactics, it is likely that certain tactics will have more effect on newcomers than others because different tactics provide information in different ways. For example, collective or formal tactics are likely to be the least important types in mediating adjustment, because both essentially depersonalize newcomers and do not allow for variability in their responses to situations. Fixed and sequential tactics are likely to be more important in this regard because they provide newcomers with content-specific information concerning their future organizational progress. However, investiture and serial socialization tactics are likely to be most important, especially in situations in which information is ambiguous or role learning is complex, because they provide the social cues and facilitation necessary during learning processes (Feldman, 1976). This argument led to the following hypothesis:

> *Hypothesis 3: Investiture and serial methods will be the most important among the six categories of socialization tactics in mediating personal adjustments to organizations.*

Self-Efficacy and Role Orientation

The introduction to this paper argued that newcomers' levels of self-efficacy may moderate the effects of socialization tactics on role orientation. Self-efficacy influences people's expectations about their abilities to perform successfully in new situations. Expectations about self-efficacy are directly related to people's perceptions of their success in dealing with past situations and to their expectations about future success. Thus, the knowledge newcomers have acquired through past experience and the strategies they have developed in the past to deal with ambiguity or uncertainty may also influence how they select information and respond to new situations.

For example, people with high levels of self-efficacy might be expected to define their new roles differently than those with low levels because of differing beliefs about their personal competency. As a result, newcomers with high levels of self-efficacy may take proactive stances toward role performance in order to demonstrate their abilities; alternatively, the socialization tactics used may not affect their subsequent role orientations. These newcomers will interpret situations as they see fit. Conversely, individuals with low levels of self-efficacy may more readily accept definitions of situations offered by others. In general, it may be argued that when organizations employ institutionalized methods of socialization, custodial role orientations are more likely to result among individuals possessing low rather than high levels of self-efficacy. This leads to the following hypothesis:

> *Hypothesis 4: A newcomer's level of self-efficacy will moderate the effects of institutionalized tactics on role orientation.*

In sum, the purposes of this study were to investigate (1) how socialization tactics affect role orientations and personal adjustments to organizations, and (2) whether level of self-efficacy moderates the effects of socialization tactics on role orientations.

METHODS

Respondents and Procedures

In order to investigate the effects of both individual and organizational variables on newcomers' adjustments to organizations, this study employed a longitudinal research design. Respondents, all M. B. A. students from two successive annual graduating classes of a major midwestern university, completed two questionnaires. They completed the first after they had accepted jobs but before they joined their organizations; they completed the second approximately five months after joining. This time period was chosen to capture newcomers' responses after they had had time to experience socialization processes and develop role orientations. Respondents received questionnaires to be completed before joining organizations at the time of their exit interviews. They were requested to mail their responses to the researchers.

The second questionnaire was mailed to respondents and follow-up letters were sent one month later.

The questionnaire completed before entry measured levels of self-efficacy. The second questionnaire measured the six socialization tactics and the six outcome variables: role orientation, role conflict, role ambiguity, job satisfaction, intention to quit, and commitment. The first questionnaire was completed by 127 respondents from a population of 282, giving a 45 percent response rate; 102 of the original 127 (an 80 percent response rate) completed the second questionnaire. Thus, overall 36 percent of the original population responded. In order to investigate whether the characteristics of the final group differed from those of the original one, the study compared the original group's M. B. A. concentrations, as expressed in their exit interviews, and the final group's functional specializations in their organizations, as expressed in the second questionnaire. In general, the proportions of respondents in various business functions like finance, marketing, and so forth, were similar to the proportions of respondents in the various M. B. A. concentrations. This provides some support for the representativeness of the final group. Respondents had joined 96 widely diverse firms located mainly in the Midwest and Southwest: major accounting firms, banks, manufacturing and service companies, as well as small, specialized organizations. There were 73 men and 29 women in the final group. The average age of respondents was 24.7 years.

Measures

Self-efficacy. Following Bandura (1977, 1978), self-efficacy was measured in terms of the people's expectations that "they can successfully execute the behavior required to produce the outcome" (Bandura, 1977: 192), in this case, mastery of role and organizational requirements. The self-efficacy scale consisted of eight items scored on 7-point Likert-type scales ranging from "strongly agree" to "strongly disagree" ($\alpha = 71$). The Appendix lists the items in this scale.

Socialization tactics. The six socialization tactics were operationally defined in terms of Van Maanen and Schein's (1979) discussion. Care was taken to ensure that each facet of each tactic figured as an item in each scale. Also, care was taken to endow questions with an active, behavioral, rather than an affective, evaluative tone in order to reduce common method variance. The Appendix lists the six 5-item scales that resulted from this procedure. Table 2 gives the reliabilities of these scales and all others as its diagonal elements. Although the different ends of the scales bear different names in Table 2 for purposes of clarity (e.g., fixed versus variable), the scales are not bipolar, but are based on a Likert-type format.

To further investigate the relationships between tactics, varimax factor analysis was performed. Table 1 gives the results of this procedure. Both of the social tactics load on Factor 1 and both of the content tactics load on Factor 2. However, formal versus informal and collective versus individual tactics load largely on different dimensions, indicating that not all formal

TABLE 1
Rotated Factor Loadings from Factor Analysis for the Socialization Tactic Scales

	Factors[a]			
Items[b]	**1**	**2**	**3**	**4**
Investiture vs. divestiture				
ID1	.57			
ID2	.74			
ID3	.61			
ID4	.62			
ID5	.76			
Serial vs. disjunctive				
SD1	.41			
SD2	.67			
SD3	.63			
SD4	.46			
SD5	.44			
Serial vs. random				
SR1		.61		
SR2		.42		
SR3		.43		
SR4		.46		
SR5		.73		
Fixed vs. variable				
FV1		.62		
FV2		.55		
FV3		.69		.45
FV4		.56		
FV5		.55		
Collective vs. individual				
CI1			.75	
CI2			.45	
CI3			.57	
CI4			.62	
CI5			.67	
Formal vs. informal				
FI1			.48	.48
FI2				
FI3				.67
FI4				.53
FI5			.46	

[a] Only loadings above .40 are presented.
[b] For items identified by abbreviations, refer to Appendix.

socialization programs were collective, nor all informal programs individual (Van Maanen & Schein, 1979). These results provide some support for the classificatory scheme in Figure 1.

Outcome measures. Role orientation was measured on a 5-item Likert-type scale using a 1-7 response format ranging from "strongly disagree" to "strongly agree." A high score on orientation indicates an innovative orientation, in which a newcomer attempts to alter procedures for performing a role, the purpose of the role itself, or both. A low score indicates a custodial orientation, in which a newcomer accepts the prescribed limits of a role in an organization. Role conflict and ambiguity were measured with the scales developed by Rizzo, House, and Lirtzman (1970).

In order to ensure that role orientation was conceptually and empirically independent of role conflict and ambiguity, a factor analysis of the 20 items composing these scales was performed. Three factors with eigenvalues greater than 1 were obtained; these accounted for 93.6 percent of the common variance. Examination of the item loadings using a .40 cutoff revealed that Factor 1 contained only those items referring to role ambiguity; Factor 2 only role conflict items; and Factor 3 only role orientation items. This supported the analytical separation of these variables.

For the personal outcome variables, commitment was measured using a shortened version of the scale by Porter, Steers, Mowday, and Boulian (1974). (See the Appendix.) Satisfaction was measured on the 7-point Faces scale (Dunham & Herman, 1975; Kunin, 1955) as a summary measure of overall job satisfaction in order to reduce method variance. Intention to quit was measured by a 2-item scale: "I will probably look for a new job in the coming year," and "I scan the newspapers and other sources for prospective jobs."

RESULTS

Intercorrelations among all variables appear in Table 2. Examining the intercorrelations between the scales measuring socialization tactics reveals a high degree of intercorrelation between five of the six scales. Investiture-divestiture has the smallest degree of intercorrelation with the other tactics ($r = .28$). However, in all cases intrascale reliability estimates (α) are of greater magnitude than interscale Pearson correlations. This suggests at least a tolerable level of discriminant validity between tactics (Buchanan, 1974). Moreover, as discussed in earlier sections of this paper, theory suggests that the use of these tactics will be highly correlated, and the pattern of correlation between the tactics supports this hypothesis. For example, investiture correlates most highly with serial socialization tactics, and sequential tactics correlate most highly with fixed tactics.

In regard to the issue of the relationship between socialization tactics and role orientation, it appears that an innovative role orientation is significantly and negatively related to each of the institutional methods of socialization. On the other hand, individualized socialization programs were associated with relatively high levels of role conflict and ambiguity plus innovative role orientations. This implies that high levels of role conflict and ambiguity are associated with performing roles in innovative ways. Either newcomers self-generate dissonance through how they perform their roles, or their actions lead others to question how they perform. For example, organizational incumbents' resisting newcomers' attempts to redefine their roles will raise the level of role conflict and ambiguity that the newcomers experience. Newcomers will be forced to trade off their desires to perform

TABLE 2
Intercorrelations among Study Variables [a,b]

Variables	x	s.d	1	2	3	4	5	6	7	8	9	10	11	12	13
1. Role orientation	20.2	7.36	(.89)												
2. Role conflict	26.6	9.16	.51	(.61)											
3. Role ambiguity	19.6	7.30	.43	.50	(.85)										
4. Commitment	47.9	11.90	-.28	-.61	-.37	(.71)									
5. Job satisfaction	5.3	1.41	-.28	-.59	-.43	.79	—								
6. Intention to quit	5.2	3.27	.33	.34	.29	-.64	-.56	(.81)							
7. Investiture vs. divestiture	26.6	5.91	-.30	-.51	-.54	.60	.65	-.37	(.79)						
8. Serial vs. disjunctive	24.8	7.03	-.53	-.47	-.59	.49	.53	-.31	.60	(.78)					
9. Fixed vs. variable	20.6	7.30	-.47	-.42	-.53	.30	.42	-.34	.41	.66	(.79)				
10. Sequential vs. random	22.6	7.55	-.52	-.44	-.51	.38	.44	-.38	.32	.70	.80	(.78)			
11. Collective vs. individual	21.3	8.54	-.35	-.07	-.28	.06	.09	-.18	.14	.57	.54	.64	(.84)		
12. Formal vs. informal	18.2	6.40	-.28	-.24	-.27	.16	.19	-.20	.03	.44	.57	.64	.70	(.68)	
13. Self-efficacy	37.9	7.33	.16	.19	-.09	-.21	-.18	.26	-.03	-.20	-.13	-.17	-.13	-.08	(.71)

[a] $N = 102$; correlation coefficients above .16 are significant at $p < .05$; those above .23, at $p < .01$.

[b] Given the nature of the scaling procedure, a positive correlation between a socialization tactic and any other variable is to be interpreted as the relationship between the institutional end of the continuum and the variable. A negative correlation indicates a relationship with the individualized end. Thus, for interpretative purposes, an active role orientation is positively correlated with divestiture, or negatively correlated with investiture. This is the case for *all* tables.

their roles innovatively against desires to reduce uncertainty by lowering levels of role conflict and ambiguity. Conversely, custodial role orientations are associated with low levels of role conflict and ambiguity, as might be expected. This issue is explored further later in the Results section.

The relationship between socialization tactics and personal outcomes also shows a consistent pattern of results. The more institutionalized the form of socialization was, the greater were expressed job satisfaction and commitment, and the lower was intention to quit. Thus, there appears to be considerable support for the hypothesized relationship between institutionalized tactics and reduction of anxiety.

Further analysis was conducted in order to investigate the relationship between personal and role outcomes and the six socialization tactics, and also to discover which tactics have most impact upon these outcomes. Since there are multiple dependent and multiple independent variables, canonical correlation analysis was used. This technique was appropriate because its aim is to develop parsimonious linear combinations of dependent and independent variables that maximize correlations between each set of linear combinations. Table 3 presents the results of the canonical analysis. Here, the structure coefficients are the beta weights used to form the respective dependent and independent linear combinations. In other words, applying

TABLE 3
Results of Canonical Analysis of the Relationships of Socialization Tactics with Personal and Role Outcomes

Canonical Roots	Eigenvalues	Wilks's Lambda	*F*	Canonical Correlations	Coefficients of Redundancy
1	2.52	.18	5.27**	.85	.51
2	0.28	.63	1.74*	.47	.64

	Canonical Set I		Canonical Set II	
Variables	Structure Coefficients	Standardized Coefficients	Structure Coefficients	Standardized Coefficients
Dependent variables				
Role orientation	-.56	-.19	.67	.92
Role conflict	-.76	-.08	-.20	-.59
Role ambiguity	-.77	-.39	.27	.31
Commitment	.78	.25	.34	.37
Job satisfaction	.86	.44	.28	.11
Intention to quit	-.51	.11	.05	.16
Independent variables				
Collective vs. individual	.26	-.41	-.77	-.91
Formal vs. informal	.33	.18	-.32	.71
Fixed vs. variable	.65	-.05	-.50	-.45
Sequential vs. random	.69	.43	-.53	-.05
Investiture vs. divestiture	.85	.55	.20	.69
Serial vs. disjunctive	.82	.38	-.43	-.30

* $p < .05$
** $p < .001$

these standardized weights to the respective variables produces the canonical variables. The standardized coefficients are the respective zero-order correlations of the variables with the linear combinations (i.e., canonical variables). Because previous research (Cooley & Lohnes, 1971) has shown that the structure coefficients have the greatest stability, they were used for interpretive purposes.

The results of the canonical analysis give two significant canonical functions, with the first accounting for most of the variance. (The coefficient of redundancy is the multivariate analog of r^2.) It is immediately obvious that the structure coefficients for both independent variable canonical functions are almost mirror images of one another. With the exception of investiture versus divestiture, whose loading falls from .85 to .20, the first canonical function represents an institutionalized pattern of socialization tactics, while the second canonical function represents an individualized pattern. Relating these results to the structure coefficients of the dependent variables reveals that for the first canonical function, high job satisfaction and commitment, and low role conflict and ambiguity load the heaviest. For the independent variables, investiture, serial, sequential, and fixed tactics load highest. The results support Hypothesis 2, stating that institutionalized tactics will be most significant in mediating personal adjustments to organizations. On the second canonical function, of the dependent variables only role orientation loads high (0.67), and individual, random, variable, and disjunctive socialization tactics, in that order, load highest. These results suggest that an innovative role orientation is the main outcome of individualized practices, or conversely, that institutionalized tactics are the tactics most likely to produce custodial role orientations. Thus, taken together, these results support Hypotheses 1 and 2 and suggest that institutional and individualized socialization tactics have different effects on role orientation and personal adjustment.

Hypothesis 3 suggested that investiture and serial tactics would be most important in mediating personal adjustments to organizations. The loadings of the six tactics on canonical function 1 support this proposition. Investiture and serial tactics load highest, followed by fixed and sequential tactics, with collective and formal tactics loading lowest. These results appear to suggest that social rather than context types of institutional socialization tactics will have more effect on newcomers' transitions into organizations.

The next issue was the effect of level of self-efficacy on role orientation. The Pearson correlation coefficients in Table 2 suggest that role orientation is positively related to self-efficacy. In order to test the hypothesis that self-efficacy moderates the effects of socialization tactics on role orientation, a series of moderated regressions were performed. A separate regression equation was run for each socialization tactic with role orientation as the dependent variable. In each equation, self-efficacy and an interaction term composed of the product of the socialization tactic and self-efficacy were included. The interaction term was significant ($p < .05$) for three of the six socialization tactics. To conserve space, this paper does not show the results of these regressions. Instead, a subgroup correlation analysis that provides similar information was subsequently performed in accordance with Peters and Champoux's (1979) recommendation. A median split on self-efficacy scores was performed and separate correlations were then computed for the relationship between role orientation and each socialization tactic. Table 4 presents correlations and subgroup differences. These results suggest that institutionalized socialization tactics result in considerably stronger custodial orientations when individuals possess low rather than high levels of self-efficacy. It appears that individuals low in self-efficacy more readily conform to definitions of situations offered by others. In other words, level of self-efficacy moderates the effects of socialization tactics on role orientation. These results provide some support for an interactionist perspective. In order to understand the nature of the newcomers' adjustments to organizations, it is necessary to consider the combined effects of individual and organizational factors.

DISCUSSION AND CONCLUSIONS

Empirical investigation of the way in which socialization tactics affect personal and role outcomes has lagged behind the theoretical conceptualization of these issues (Feldman, 1976). In recent work, both Van Maanen and Schein (1979) and Louis (1980) suggested that reducing uncertainty is the major goal of newcomers to organizations and that the methods of socialization that organizations use can influence the way newcomers respond to their organizations. The findings of this study support this hypothesis, and the canonical results reported here reveal a pattern of relationships between tactics and outcomes supporting the proposition that different socialization tactics lead to different outcomes of socialization. The social dimensions of

TABLE 4
Moderating Influence of Self-Efficacy on Correlations between Role Orientation and Socialization Tactics

	Correlations with Role Orientation		
Socialization Tactics	Low Self-Efficacy	High Self-Efficacy	Significance of Difference[a]
Collective vs. individual	-.45***	-.25*	.10
Formal vs. informal	-.42***	-.12	.06
Sequential vs. random	-.57***	-.43***	n.s.
Fixed vs. variable	-.63***	-.22	.01
Investiture vs. divestiture	-.48***	-.09	.05
Serial vs. disjunctive	-.68***	-.33**	.05

[a] The significance of the difference between subgroup correlations was tested applying Fisher's r to z transformation and examining the differences using a Z test.

* $p < .05$

** $p < .01$

*** $p < .001$

socialization—investiture and serial processes—appear to be particularly significant in influencing role orientations and subsequent adjustments to organizations. However, the content of the information given to newcomers concerning role requirements and their future organizational progress also appears to reduce the uncertainty surrounding the entry process. Both collective and formal socialization tactics were relatively less influential, perhaps because they define newcomers as learners and increase the psychological distance between newcomers and organizations. The importance the literature has attributed to these variables may be a consequence of authors' using highly unusual or extreme settings, such as military socialization programs, as illustrations. In such settings recruits have no other choice but to respond to organizational commands. Very few organizations have such captive audiences to experiment on.

These results also suggest that the different ends of the six continua of socialization tactics, rather than a mix of tactics, are instrumental in leading to particular role orientations. Moreover, contrary to Van Maanen and Schein's (1979) argument, both fixed and investiture tactics, rather than variable and divestiture tactics, are associated with custodial responses. This pattern supports this study's hypothesis and suggests that two factors combine to discourage newcomers from performing their roles in ways that are contrary to those customary in an organization. These discouraging factors are ability to predict progress in an organization and definitions of situations offered by others. The finding that role orientation is positively correlated with role conflict and ambiguity further supports this interpretation; failure to accept significant others' definitions of situations results in increasing uncertainty or anxiety for newcomers.

As argued in an earlier section, newcomers are likely to have proclivities to interpret situations in certain ways and to respond to them in their own terms. This study found that newcomers' levels of self-efficacy moderated the effects of institutionalized tactics on role orientation, in support of this hypothesis. This finding suggests that newcomers use self, as well as others, as a locus of control when evaluating situations (Weiss, 1978). The results suggest that newcomers high in self-efficacy will tend to define situations themselves even when their roles or progressions in organizations are prescribed.

However, caution should be exercised in interpreting these results because this study relied exclusively on self-reported measures. Studies employing only self-reported data have raised concern among organizational researchers, because common method variance may affect the magnitude of relationships between work characteristics and work outcomes. The use of a longitudinal research design partially obviated this problem so far as the moderating effects of self-efficacy are concerned because self-efficacy measured at time 1 was analyzed with tactics and outcomes measured at time 2. Moreover, as discussed in the Methods section, care was taken to operationally define the measures of socialization tactics in as objective a way as possible, and the relationships between the outcome measures are consistent with those found in many studies. Further, all hypotheses were directional and all were supported as stated. Nevertheless, it would have been useful to have had independent sources like supervisors confirm the nature of socialization tactics. The original research design attempted to get such confirmation, but respondents' anxieties about anonymity and feedback to supervisors prevented its use. The procedure in effect increased newcomer anxiety. Finally, all methodologies have costs and benefits. Although common method variance may be a cost of survey research, this study gained one main benefit from its use. This benefit was the opportunity to examine socialization practices across widely diverse organizations, which provided the variability in responses necessary for meaningful analysis.

In conclusion, it is suggested that future research be conducted across different firms that hire similar candidates but treat them differently. These settings should produce different results that will allow further investigation of the effects of socialization tactics on newcomers' adjustments. It would be interesting to study organizations' socialization practices in conjunction with newcomers' tasks. There is likely to be some correspondence between the demands of tasks and socialization practices; this correspondence would resemble in kind the relationships between choice of leader behaviors and task constraints that path-goal theories of leadership have posited (House, 1971). Such research may make it possible to fathom why firms use certain kinds of tactics rather than others. It would also be interesting to investigate how organizations' sizes or natures influence their choices of socialization tactics. It is likely that research that considers the requirements of different jobs as to skills and roles will find factors based on individual differences even more salient than did this study. Finally, also deserving of investigation is the degree to which newcomers may, because of individual differences, self-select organizations because they know what kinds of socialization and training the organizations use. This issue is likely to be particularly important in the case of newcomers who join unusual socialization programs such as the military's. The measures of socialization and self-efficacy developed in this study may, with suitable modification, prove useful in such future research.

REFERENCES

Bandura, A. 1977. Self-efficacy: Towards a unifying theory of behavioral change. ***Psychological Review***, 84: 191–215.

Bandura, A. 1978. The self-system in reciprocal determination. ***American Psychologist***, 33: 344–358.

Buchanan, B. 1974. Building organizational commitment: The socialization of managers in work organizations. ***Administrative Science Quarterly***, 19: 533–546.

Cooley, W., & Lohnes, P. R. 1971. ***Multivariate data analysis***. New York: John Wiley & Sons.

Dunham, R., & Herman, J. 1975. Development of female Faces scale for measuring job satisfaction. ***Journal of Applied Psychology***, 60: 629–631.

Feldman, D. C. 1976. A contingency theory of socialization. ***Administrative Science Quarterly***, 21: 433–452.

House, R. J. 1971. A path-goal theory of leader effectiveness. ***Administrative Science Quarterly***, 16: 321–339.

Hughes, E. C. 1958. ***Men and their work***. Glencoe, Ill.: Free Press.

Jones, G. R. 1983a. Psychological orientation and the process of organizational socialization: An interactionist perspective. ***Academy of Management Review***, 8: 464–474.

Jones, G. R. 1983b. Organizational socialization as information processing activity: A life history analysis. ***Human Organization***, 42: 314–320.

Katz, R. 1978. Job longevity as a situational factor in job satisfaction. ***Administrative Science Quarterly***, 23: 205–223.

Kunin, T. 1955. The construction of a new type of attitude measure. ***Personnel Psychology***, 8: 65–78.

Louis, M. R. 1980. Surprise and sense-making: What newcomers experience in entering unfamiliar organizational settings. ***Administrative Science Quarterly***, 25: 226–251.

Peters, W. S., & Champoux, J. E. 1979. The use of moderated regression in job redesign decisions. ***Decision Sciences***, 10: 85–95.

Porter, L., Steers, R. M., Mowday, R. T., & Boulian, P. V. 1974. Organizational commitment, job satisfaction, and turnover amongst psychiatric technicians. ***Journal of Applied Psychology***, 59: 603–609.

Rizzo, J. R., House, J. H. & Lirtzman, S. I. 1970. Role conflict and ambiguity in complex organizations. ***Administrative Science Quarterly***, 15: 150–163.

Schein, E.H. 1968. Organizational socialization and the profession of management. ***Industrial Management Review***, 2: 59–77.

Van Maanen, J. 1977. Experiencing organizations. In J. Van Maanen (Ed.), ***Organizational careers: Some new perspectives***: 15–45. New York: John Wiley & Sons.

Van Maanen, J., & Schein, E. H. 1979. Towards a theory of organizational socialization. In B. M. Staw (Ed.), ***Research in Organizational Behavior***, vol. 1: 209–264. Greenwich, Conn.: JAI Press.

Weiss, H. 1978. Social learning of work values in organizations. ***Journal of Applied Psychology***, 63: 711–718.

APPENDIX

Unless otherwise noted, responses were measured on 7-point scales ranging from "strongly disagree" to "strongly agree." (R) indicates reverse scoring.

Scales Measuring Socialization Tactics

Collective versus individual:

CI1 In the last six months, I have been extensively involved with other new recruits in common, job related training activities.
CI2 Other newcomers have been instrumental in helping me to understand my job requirements.
CI3 This organization puts all newcomers through the same set of learning experiences.
CI4 Most of my training has been carried out apart from other newcomers. (R)
CI5 There is a sense of "being in the same boat" amongst newcomers in this organization.

Formal versus informal:

FI1 I have been through a set of training experiences which are specifically designed to give newcomers a thorough knowledge of job related skills.
FI2 During my training for this job I was normally physically apart from regular organizational members.
FI3 I did not perform any of my normal job responsibilities until I was thoroughly familiar with departmental procedures and work methods.
FI4 Much of my job knowledge has been acquired informally on a trial and error basis. (R)
FI5 I have been very aware that I am seen as "learning the ropes" in this organization.

Investiture versus divestiture:

ID1 I have been made to feel that my skills and abilities are very important in this organization.
ID2 Almost all of my colleagues have been supportive of me personally.
ID3 I have had to change my attitudes and values to be accepted in this organization. (R)
ID4 My colleagues have gone out of their way to help me adjust to this organization.
ID5 I feel that experienced organizational members have held me at a distance until I conform to their expectations. (R)

Sequential versus random:

SR1 There is a clear pattern in the way one role leads to another or one job assignment leads to another in this organization.
SR2 Each stage of the training process has, and will, expand and build upon the job knowledge gained during the preceeding stages of the process.
SR3 The movement from role to role and function to function to build up experience and a track record is very apparent in this organization.
SR4 This organization does not put newcomers through an identifiable sequence of learning experiences. (R)
SR5 The steps in the career ladder are clearly specified in this organization.

Serial versus disjunctive:

SD1 Experienced organizational members see advising or training newcomers as one of their main job responsibilities in this organization.
SD2 I am gaining a clear understanding of my role in this organization from observing my senior colleagues.
SD3 I have received little guidance from experienced organizational members as to how I should perform my job. (R)
SD4 I have little or no access to people who have previously performed my role in this organization. (R)
SD5 I have been generally left alone to discover what my role should be in this organization. (R)

Fixed versus variable:

FV1 I can predict my future career path in this organization by observing other people's experiences.
FV2 I have a good knowledge of the time it will take me to go through the various stages of the training process in this organization.
FV3 The way in which my progress through this organization will follow a fixed timetable of events has been clearly communicated to me.
FV4 I have little idea when to expect a new job assignment or training exercise in this organization. (R)

FV5 Most of my knowledge of what may happen to me in the future comes informally, through the grapevine, rather than through regular organizational channels. (R)

Scale Measuring Self-Efficacy

My new job is well within the scope of my abilities.
I do not anticipate any problems in adjusting to work in this organization.
I feel I am overqualified for the job I will be doing.
I have all the technical knowledge I need to deal with my new job, all I need now is practical experience.
I feel confident that my skills and abilities equal or exceed those of my future colleagues.
My past experiences and accomplishments increase my confidence that I will be able to perform successfully in this organization.
I could have handled a more challenging job than the one I will be doing.
Professionally speaking, my new job exactly satisfies my expectations of myself. (R)

Scales Measuring Outcomes

Role orientation:

Responses were measured on a 7-point scale ranging from "very false" to "very true."
I have made an attempt to redefine my role and change what I am required to do.
While I am satisfied with my overall job responsibilities, I have altered the procedures for doing by job.
I have changed the mission or purpose of my role.
The procedures for performing my job are generally appropriate in my view. (R)
I have tried to change the procedures for doing my job and to institute new work goals.

Commitment to organization:

I talk up this organization to my friends as a great organization to work for.
I feel very little loyalty to this organization. (R)
I would accept almost any type of job assignment in order to keep working for this organization.
I am proud to tell others that I am a part of this organization.
I could just as well be working for a different organization as long as the type of work was similar. (R)
It would take very little change in my present circumstances to cause me to leave this organization. (R)
I am extremely glad I chose this organization to work for over others I was considering at the time I joined.
There is not much to be gained by sticking with this organization indefinitely. (R)
For me, this is the best of all possible organizations for which to work.
Deciding to work for this organization was a definite mistake on my part. (R)

All other items and response formats used were identical to those of the established scales referred to in the Methods section.

Gareth R. Jones received his Ph.D. degree in organizational theory from the University of Lancaster, U.K.; he is Assistant Professor of Management at Texas A & M University. His current research interests center on organizational control and the application of transaction costs analysis to organization-environment relations and strategy formulation.

© Academy of Management Journal
1986, Vol. 29, No. 2, 280–304.

OUTCOMES OF AUTONOMOUS WORKGROUPS: A LONG-TERM FIELD EXPERIMENT

TOBY D. WALL
NIGEL J. KEMP
PAUL R. JACKSON
CHRIS W. CLEGG
University of Sheffield, England

A quasi-experimental design was used to study the long-term effects of implementation of autonomous workgroups in a manufacturing environment. Results indicated a substantial and lasting effect on employees' intrinsic job satisfaction, a more temporary effect on extrinsic job satisfaction, and no consequences for work motivation or performance. At an organizational level, however, improvements in productivity were made possible through elimination of supervisory positions; and contrary to prediction, labor turnover increased.

Autonomous workgroups represent a form of work design that has evolved from the general sociotechnical systems approach to organizational analysis and design (Emery, 1959; Katz & Kahn, 1978; Susman, 1976). Although comparisons are difficult, it seems that more organizations have implemented work designs with autonomous workgroups on their shopfloors than have implemented job characteristics approaches oriented to individuals. Certainly, many observers have commented on the increased use of autonomous workgroups in a number of industrialized countries (Cummings, 1978; Pasmore, Francis, Haldeman, & Shani, 1983). Paradoxically, empirical evidence concerning the effects of such groups on employee attitudes and behavior has not shown a corresponding advance. The present study, a large scale, long-term, outcome evaluation of autonomous workgroups based on a quasi-experimental research design, helps to remedy this deficiency. By way of introduction we outline the nature of autonomous workgroups, consider their theoretically predicted effects on employee attitudes and behavior, and discuss the need for the kind of empirical investigation described in this paper.

THEORETICAL BACKGROUND AND RESEARCH AIMS

The Nature of Autonomous Workgroups

The key feature of autonomous workgroups is a high degree of self-determination by employees in the management of their day-to-day work. Typically this includes collective control over the pace of work, distribution of tasks, organization of breaks, and collective participation in the recruitment and training of new members (Gulowsen, 1972). Often direct supervision is unnecessary (Emery, 1980). To allow for such self-management, the organization of work commonly has these three characteristics: (1) employees with functionally interrelated tasks who collectively are responsible for end products; (2) individuals who have a variety of skills so they may undertake all or a large proportion of groups' tasks; and (3) feedback and evaluation in terms of the performance of whole groups.

It will not go unnoticed that the properties prescribed for autonomous workgroups parallel the hallmarks of the job characteristics approach to job design (Hackman & Lawler, 1971; Hackman & Oldham, 1976). These are variety, task identity, autonomy, and feedback. On this basis, Rousseau (1977) argued for a synthesis of the two approaches, as did Cummings (1978) and Denison (1982). Hackman (1977) went further and suggested that autonomous group working is the more powerful of the two approaches because groups can undertake much larger pieces of work than individuals can and so allow more fundamental manipulation of work characteristics.

Thus, previous work has specified the core characteristics of autonomous workgroups relatively clearly, tying them not only to the sociotechnical literature, but also to job design theory in general. The main distinction between autonomous workgroups and the job characteristics approach concerns level of application and analysis, not content. For autonomous groups, job characteristics are a property of group work, not of individual jobs. However, this difference is not apparent in the predicted effects of these two approaches to work design.

Effects of Implementation of Autonomous Workgroups

Sociotechnical theorists have explicitly, or more often implicitly, described rationales for and effects of autonomous workgroups in terms of individual psychological reactions (Davis, 1966; Kelly, 1978). The underlying assumptions have been that this way of organizing work is intrinsically motivating and enhances employee satisfaction; from these reactions follow improved group performance and reduced labor turnover. Here again there is a direct correspondence to the individual job design approach; in their job characteristics model, Hackman and Oldham (1976) posited these same causal relationships. It is thus not surprising that Hackman (1983) developed a normative theory of workteam effectiveness from the literatures on job design and small group behavior. Though it deals with wider issues concerning the composition of workgroups, at the heart of this integrative model are the same propositions concerning the implications of autonomous group work for motivation, member satisfaction, performance, and labor turnover.

The literature suggests two further effects of autonomous workgroups, increased organizational commitment (Emery, 1959) and improved mental health (Herbst, 1974). Indeed, Herbst reported that the "starting point of the early coal studies was provided by psychiatric investigations . . . which gave evidence of an epidemic incidence of psychosomatic disorders among miners" (1974: 3). The implication is that a change to autonomous group work would reduce such disorders. Other research (Alcalay & Pasick, 1983; Broadbent, 1985; Fraser, 1947; Gardell, 1971; Herzberg, 1966; Karasek, 1978; Kornhauser, 1965; Wall, Clegg, & Jackson, 1978; Wall & Clegg, 1981) undertaken in many different work settings gives general support to this view. Neither organizational commitment nor mental health is a central element in existing theoretical propositions; nevertheless, these outcome variables were included in the present investigation to continue that line of research.

Need for Field Experimentation

Although the nature and predicted effects of autonomous workgroups are relatively clear from the theoretical literature, research on their effectiveness has been less definitive. Lawler concluded that "available research consists mainly of poorly documented case studies and correlational studies" (1977: 576) and called for research based on quasi-experimental designs. Cummings, Molloy, and Glen reviewed 58 studies with work experiments and found "the studies fared badly when assessed against the [quasi-experimental] criteria used in this critique" (1977: 702). Pasmore and colleagues argued that "more closely controlled experimental introductions of work restructuring are called for, with appropriate groups serving as controls during the experiments" (1983: 1198); Roberts and Glick (1981) endorsed this view in relation to job design research in general.

With respect to studies of autonomous workgroups, Lawler's (1977) allegation of low quality was perhaps harsh. There exist several careful accounts of such innovations (Davis & Valfer, 1965; Kemp, Clegg, & Wall, 1980; Kemp, Wall, Clegg, & Cordery, 1983; Rice, 1958; Trist & Bamforth, 1951; Trist, Susman, & Brown, 1977). Nevertheless, it is a valid point that very few of the findings are based on research designs allowing causal inferences. A point less often raised, but nonetheless important, concerns the time spans of research into work design. Most studies, particularly those with relatively rigorous designs, limit their evaluation to but a few months. But work designs try to achieve enduring changes in employees' attitudes and behaviors, so a perspective in terms of years rather than months is desirable. Previous research can be largely interpreted as supporting only short-term and thus possibly transitory effects (Wall & Clegg, 1981).

The absence of long-term experimental studies of autonomous workgroups is not altogether surprising. Research of this nature is time consuming and demands considerable resources. Moreover, the opportunities are few. Because this research is of immense theoretical and practical importance, such opportunities should be exploited when they arise. Such an opportunity presented itself to us, as the remainder of this paper describes.

Aims of the Present Study

The primary aim of this long-term experiment was to examine the effects of autonomous workgroups on employees' attitudes and behaviors. More specifically, this experiment tested the proposition that this form of work design enhances work motivation, job satisfaction, and group performance, and reduces labor turnover. We also included organizational commitment and mental health as outcome variables for exploratory purposes. The study's secondary aim was to report and comment on the organizational issues arising from extended use of autonomous workgroups.

METHODS

Background, Setting, and the New Form of Work Organization

This study was set within a large, nonunionized, British company that produces confectionery at several sites for the home market and export. When planning a new factory in the southeast of England, senior management decided to take the opportunity to try a new method of working that recognized the needs and capitalized upon the skills of their employees. This choice reflected an ideological commitment to increasing responsibility on the shopfloor and an understanding of modern principles of social science governing work design. The company engaged a consultant from the Tavistock Institute of Human Relations, Dr. Lisl Klein, to work alongside managers, engineers, and architects throughout the planning process. The outcome of this collaboration was a modern, spacious, and comfortable factory whose management structures and practices were designed to support a well-developed form of autonomous group working.

From the outset of the study we were independent from the consultant and were not involved in the design and implementation processes. Our role was to evaluate the effects of autonomous workgroups over a period of three years, beginning with the start-up of production. We accepted no fees from the organization and guaranteed both confidentiality of individual views and company anonymity.

The intent of the new form of work organization was to provide shopfloor employees with substantial autonomy in carrying out day-to-day production along with a high level of involvement in operational decision making. In practice, production employees worked in groups of 8 to 12 people, all of whom were expected to carry out each of eight types of jobs involved in the production process. Group members were collectively responsible for allocating jobs among themselves, reaching production targets and meeting quality and hygiene standards, solving local production problems, recording production data for information systems, organizing breaks, ordering and collecting raw materials and delivering finished goods to stores, calling for engineering support, and training new recruits. They also participated in selecting new employees. Within each group, individuals had considerable control over the amount of variety they experienced by rotating their tasks, and each production group was responsible for one product line. Group members interacted informally throughout the working day but made the most important decisions—for example, regarding job allocation—at formal weekly group meetings where performance was also discussed. This method of working thus meets all the major criteria of autonomous group work.

However, perhaps the most marked feature of this form of work organization concerned supervision. Since nearly all the functions typically carried out by supervisors were the responsibility of shopfloor employees, the position of supervisor did not exist. Groups reported directly to a support manager who had overall responsiblity for some 40 to 70 employees and whose role was to provide the necessary support and guidance for the development of both social and technical skills within the workgroups. Support managers were responsible for providing the facilities and creating the conditions within which groups could operate in a semiautonomous fashion, while maintaining disciplinary procedures and running an effective production facility.

With one exception, the managers employed in the new factory were appointed from elsewhere in the company. Some were highly committed to the new working method, some open-minded, and others somewhat skeptical. None had previous experience of group working nor of developing a greenfield site.[1] They began work on site prior to the factory's commission. Several were involved in planning the factory over a two-year period. With help from the factory personnel manager, managers were responsible for recruiting employees for their own departments. Inevitably they attempted to select and recruit people who would be able to work in groups, but they had no specific operational criteria to guide this process. Immediately prior to the start-up of production, all employees took a one-week induction course run by the factory management; the course focused on how the factory would be organized and managed with particular reference to the theory and practice of autonomous group work. Such induction courses continue to be run on a smaller scale than the original effort for all new employees. Over time, management transferred key personnel, those most skilled and experienced in the production process, within the factory to help develop new workgroups. In short, the workgroups have not been stable over time. In the factory as a whole, several features reflected the philosophy underlying autonomous group working: a single-status canteen, a single-status car park, no timeclocks, monthly payment by credit transfer, and a flat (three-level) management structure.

Research Design

The overall design of this study was quasi-experimental. It consisted of four conditions and up to three measurement occasions, as Table 1 shows. Condition A, the main experimental group, included all day-shift employees in the new factory who from the outset of production worked under the system of autonomous workgroups described in the preceding section. The company initially recruited 42 employees; the number rose to 84 during the course of the study. The numbers of employees reported in Table 1 correspond to the actual numbers of employees participating in the study.

---

[1] The term "greenfield" is used here, and subsequently, to refer to a site where a new factory is built and commissioned by the parent company. It is thus specifically designed for current purposes, unlike existing premises, which were previously used for other purposes, and then refurbished for new ones.

Within the overall research design, condition B served initially as a nonequivalent control group and later as a second experimental group. It included all employees on the evening shift at the new factory who made the same products using the same machinery as their day-shift (condition A) counterparts. For the first 12 months of the investigation their work was conventionally organized. Employees were allocated to particular tasks and a supervisor who reported to a shift manager had control over all aspects of production. Employees had none of the discretionary areas of responsibility afforded to those in condition A. After 12 months of production at the new factory, however, the company decided to extend autonomous group work to this evening shift, largely at the request of the employees themselves. At this point condition B became a second experimental group. This group initially had 25 individuals and grew to 60 by the end of the study. At the point of change to autonomous group work, the supervisor was promoted to support manager and the original manager transferred elsewhere in the company.

There were two reasons for the conventional organization of the evening shift at the outset of production. The first was to employ a supervisor and manager from a sister factory that was closing. The second was to avoid the risk of introducing autonomous group work with inexperienced employees whose schedule left them remote from the normal managerial support structure.

Conditions C and D were two further nonequivalent control groups, representing day and evening shifts in a factory elsewhere in the organization where jobs had conventional designs. Supervisors allocated particular tasks to employees and took responsibility for operational decision making. The products and technology were very similar to those in the greenfield site, the same overall company policy was in force, and pay levels were equivalent. Condition C initially included 290 employees; during the course of the study the number available dropped to 227. Condition D had 64 employees.

All employees involved in the study received payment based on standard rates reflecting an overall company policy of remunerating its employees at a little above the median rate for manufacturing companies in the area. There was no payment-by-results or group or company bonus scheme. However, overtime was available at an increased hourly rate of pay.

As originally designed, the study involved three waves of measurement for conditions A and B, with measures after 6, 18, and 30 months of production, and two waves of measurement for conditions C and D at times corresponding to the first two collections of data for conditions A and B. The primary focus for the company was a full-scale evaluation of the greenfield site, so the three measurements there were a matter of course. We negotiated the two measurements at the conventional site to enable

TABLE 1
Research Design, Cell Sizes, and Predicted Patterns of Results

	Condition A	Condition B	Condition C	Condition D	
Measurement Occasions	Greenfield Site, Day Shift	Greenfield Site, Evening Shift	Established Site, Day Shift	Established Site, Evening Shift	Predicted Patterns of Results
Time 1 (6 months)					
Work design	Autonomous	Traditional	Traditional	Traditional	
Cell size	N = 37	N = 24	N = 233	N = 54	A1 > B1, C1, D1
Group/time notation	A1	B1	C1	D1	
Time 2 (18 months)					
Work design	Autonomous	Autonomous	Traditional		
Cell size	N = 45	N = 42	N = 159	—	A2, B2 > C2
Group/time notation	A2	B2	C2		
Time 3 (30 months)					
Work design	Autonomous	Autonomous			
Cell size	N = 76	N = 50	—	—	A3, B3
Group/time notation	A3	B3			
Predicted patterns of results	A1, A2, A3	B1 < B2, B3	C1, C2		

comparisons. In practice it was impossible to complete the final data collection for condition D as the evening shift at the established factory was discontinued. The response rate was high at both sites throughout the study, varying from 70 to 95 percent of the total workforce.

As for predictions based on the overall research design, the expected pattern of results was clear. A simple effect for work design would show a pattern corresponding directly with the existence of autonomous workgroups and unaffected by other factors such as site, shift, or time. Let us assume for the sake of argument that autonomous workgroups would have a positive effect on job satisfaction. Then employees in conditions A1, A2, A3, B2, and B3 should record higher levels of job satisfaction than their counterparts in conditions B1, C1, C2, and D1. Within this global comparison, however, more precise predictions are embedded. For the first time period, for example, the predicted pattern would be of the form A1 > B1, C1, D1. Lack of difference between conditions C1 and D1 and lack of difference between conditions B1 and D1 would help to discount shift and site effects, respectively. This would strengthen the interpretation that if condition A1 differs, the effect is attributable to autonomous group work. Taking the prediction over time, it is clear that condition B should change between the first and second measurement occasions (B1 < B2, B3), whereas conditions A and C should remain constant, but at different levels, since condition A scores will be augmented from the outset by autonomous group work. The final column and row of Table 1 shows the overall pattern of results predicted if work design rather than site, shift, or time has significant effects.

This design avoided the threats to internal validity presented by history, maturation, testing, instrumentation, and statistical regression (Campbell & Stanley, 1966; Cook & Campbell, 1979a). To some extent it reduced the threats of "compensatory rivalry" and "resentful demoralization" (Cook & Campbell, 1979b: 81–82) because employees in conditions C and D were so distant from A and B—some 250 miles—as to allow little awareness of differences in work designs. Selection and its interaction with other threats and mortality posed the major potential challenges to internal validity. We discuss these later under Supplementary Analyses.

Thus the overall research design had strong internal validity and potentially high external validity. It should be noted that we could not deploy the full design for the behavioral variables. During the first six months at the greenfield site, the reliability of the newly commissioned machinery was substantially inferior to that at the established site. This had both direct and indirect effects on operators' rates of productivity. To separate human performance from machine-caused production losses proved impossible, and performance indices at this stage were consequently unreliable. Similarly, given the small numbers in the groups at the greenfield site during the first six months, the incidence of labor turnover was too small to justify separate analysis. As the period of assessment lengthened and the numbers of employees increased, a stable pattern emerged. Thus we assessed the relationships of autonomous group work to these behavioral measures by gross comparisons between the two sets of conditions; the Results section considers these relationships separately.

Data Collection and Analyses

Data on employees' perceptions and attitudes were obtained by questionnaires administered after each of the specified time periods. Confidentiality was assured, and completion during work time was voluntary. We obtained figures on labor turnover and performance from factory records. These indices are described in greater detail under Results. Throughout the study we maintained regular contact with the sites and made periodic visits to observe work in progress and interview both shopfloor employees and managers.

Self-report measures. For use as a manipulation check we developed a scale to measure perceived autonomy of workgroups. This measure included 13 items based on the criteria offered by Gulowsen (1972) for designing autonomous workgroups. Thus, we asked respondents the extent to which their workgroups got involved in specific activities of which these are examples: organizing time for breaks, allocating jobs amongst yourselves, setting the pace of work, and solving local problems. Responses were measured on 5-point formulas. Scale reliabilities (α) at each of the three measurement occasions were $\alpha = 0.82, 0.79, 0.84$.

Work attitudes were measured on scales developed and standardized in previous research on blue-collar employees (Clegg & Wall, 1981; Cook & Wall, 1980; Cook, Hepworth, Wall, & Warr, 1981; Warr, Cook, & Wall, 1979). These were: intrinsic job satisfaction ($\alpha = 0.82, 0.85, 0.79$);[2] extrinsic job satisfaction ($\alpha = 0.70, 0.74, 0.75$); intrinsic job motivation ($\alpha = 0.87, 0.90, 0.83$); and organizational commitment ($\alpha = 0.77, 0.70, 0.74$). Appendix A gives numbers of items in each scale and examples of items.

The General Health Questionnaire (Goldberg, 1972) concluded the questionnaire. This self-administered screening test designed to detect diagnosable, nonpsychotic, psychiatric disorders measures important aspects of mental health. Previous work has established the reliability and validity of this measure for general populations (Goldberg, 1978). The present study used the 12-item version, which has been shown to be both reliable and valid for employed samples (Banks, Clegg, Jackson, Kemp, Stafford, & Wall, 1980; Banks & Jackson, 1982; Jackson, Stafford, Banks, & Warr, 1983; Wall & Clegg, 1981). This questionnaire asks people if they have recently lost much sleep over worry, felt constantly under strain, been able to enjoy their normal day-to-day activities, been losing confidence in themselves, and so forth. Items have 4-point response formats ranging from "not at all" to "much more than usual." We computed overall scale scores for the three measurement occasions ($\alpha = .84, .83, .94$).

Workforce characteristics. The study is based on data from a total of 545 employees, 83 percent of whom were women. Their average age was 31 years; average age differed between shifts, evening-shift workers being older, but not between sites. Length of service naturally differed between the sites; it averaged 1.5 years at the end of the investigation for those in the new factory and had a mean of 5.7 years for those in the established factory.

---

[2] Reliabilities here and subsequently are given for each of the three measurement occasions.

Statistical analyses. Statistical analysis proceeded in two main stages. First, we applied generalized least-squares analysis to the total data set in order to provide an omnibus test of the study's predictions. Second, we examined more specific predicted patterns of results within the total data set using (1) one-way analysis of variance across conditions and (2) one-way analysis of variance over time for both independent groups and repeated measures. Appendix B describes our rationale for his dual approach.

RESULTS

The findings are presented in three parts: (1) results from the primary analyses on the perceptual and affective outcome data, (2) supplementary analyses examining the viability of rival explanations for the above results, and (3) findings related to the performance, labor turnover, and wider organizational consequences of autonomous workgroups.

Perceived Work Characteristics, Work Attitudes, and Mental Health

The initial analyses defined a regression equation for each attitudinal measure as the dependent variable, with shift (day versus evening), site (conventional versus greenfield), time (three measurement occasions), and work design (presence versus absence of autonomous workgroups) as predictors. We derived parameter estimates by generalized least squares (see Appendix B).

Throughout the analyses each dependent variable was examined separately. Such an approach may lead to misleading inferences when several dependent variables are highly related. In the present study, however, this was not the case, as is demonstrated in Table 2. This shows the correlations between the major variables in the study for the first time period only since the results were stable over time. Furthermore, dependent variables showed a diversity of response. These factors support the use of parallel analysis of each variable ignoring other dependent variables.

TABLE 2
Intercorrelations among Dependent Variables at Time 1[a]

Dependent Variables	1	2	3	4	5
1. Intrinsic job satisfaction	–				
2. Extrinsic job satisfaction	.77	–			
3. Intrinsic job motivation	.30	.42	–		
4. Organizational commitment	.40	.52	.64	–	
5. Mental health[b]	−.27	−.30	−.12	−.30	–

[a] Sample size varied betweeen 271 and 344.

[b] Lower scores signify better mental health.

With respect to perceived autonomy of workgroups, employees under autonomous work conditions gave significantly higher responses than their counterparts experiencing conventional work designs ($\beta = 0.68, p < .001$). This finding served as a manipulation check, confirming that employees perceived actual differences in work designs. Moreover, the size of this effect illustrates substantial differences between the two sets of work conditions, which thus allows a realistic test of the attitudinal effects of this form of work design.

Work design also had effects on intrinsic job satisfaction ($\beta = 0.34$, change in R^2 of 2%, $p < .001$); extrinsic job satisfaction ($\beta = 0.30$, change in R^2 of 1%, $p < .01$); and organizational commitment ($\beta = 0.25$, change in R^2 of 1%, $p < .01$); but there was no effect for intrinsic job motivation ($\beta = 0.16$, n.s.) or mental health ($\beta = -0.04$, n.s.). Thus, this overall analysis supported three of the five predictions concerning individual attitudinal effects.

These analyses provided an overall test of the effects of autonomous workgroups, but they do not suffice to test the specific patterns of findings that the research design allows and causal interpretation requires. Initially, these patterns concern differences between groups at each measurement time (Table 1, final column). Table 3 shows relevant results. These results show that for perceived autonomy of workgroups we obtained the specific predicted pattern of results. On the first measurement occasion, employees in condition A1 recorded substantially higher scores than the three control groups, which formed a homogeneous set. Similarly, on the second measurement occasion, when employees in condition B2 had changed to autonomous group work, their scores matched those for condition A2 and differed from scores for condition C2. The only minor anomaly is at the third measurement, where a difference in mean scores exists between conditions A3 and B3, though both scores are high. These findings confirm those of the initial regression analysis and moreover demonstrate the full expected pattern of results. Thus, the manifest changes in work design clearly took effect and can be seen to persist for the duration of the study.

To the extent that work design determines work attitudes and mental health, the patterns of scores recorded above should be duplicated for the affective measures. Such is the case for intrinsic job satisfaction. The differences across conditions on the first and second measurement occasions showed the predicted patterns, and on the final measurement occasion the conditions did not differ. For extrinsic job satisfaction we obtained an equivalent set of findings.

However, for the remaining three dependent variables differences across conditions do not correspond to the specific patterns required to indicate an effect for work design. Results for intrinsic job motivation and mental health confirm those of the initial regression analysis. But although the first analysis suggested that work design had an effect on organizational commitment, the second led us to reject such an interpretation. There were clear instances in which the levels of organizational commitment under traditional forms of work design matched those recorded for autonomous workgroups. Thus we

TABLE 3
Results of Analyses of Variance for Independent Measures across Groups

Dependent Variables	Measurement Occasions[a]	Condition A[b] Means	Condition B[c] Means	Condition C[d] Means	Condition D[e] Means	F	df	Tukey Contrast Test
Perceived workgroup autonomy	Time 1	44.0	26.2	31.6	31.5	16.8**	3,247	A1 > B1, C1, D1
	Time 2	45.6	42.4	31.8		50.2**	2,194	A2, B2 > C2
	Time 3	47.0	38.7			22.1**	1,101	A3 > B3
Intrinsic job satisfaction	Time 1	37.8	31.6	29.5	31.6	12.6**	3,302	A1 > B1, C1, D1
	Time 2	36.2	34.6	29.7		19.6**	2,224	A2, B2 > C2
	Time 3	36.3	35.4			<1.0	1,113	–
Extrinsic job satisfaction	Time 1	46.8	41.1	37.5	40.9	18.0**	3,302	A1 > B1, C1, D1
	Time 2	44.0	44.7	38.1		21.5**	2,224	A2, B2 > C2
	Time 3	43.5	42.3			<1.0	1,113	–
Intrinsic job motivation	Time 1	35.2	36.4	31.3	36.5	11.5**	3,307	A1, B1, D1 > C1
	Time 2	34.0	35.9	32.3		5.0*	2,228	A2, B2 > C2
	Time 3	34.3	35.9			3.0	1,116	–
Organizational commitment	Time 1	50.7	48.6	41.3	47.5	18.9**	3,315	A1, B1, D1 > C1
	Time 2	48.1	51.1	41.4		24.7**	2,225	A2, B2 > C2
	Time 3	46.3	48.3			2.0	1,113	–
Mental health[f]	Time 1	18.8	22.2	22.1	19.7	8.3**	3,314	B1, C1 > A1, D1
	Time 2	19.6	20.9	21.4		2.8	2,220	–
	Time 3	21.0	21.7			<1.0	1,114	–

[a] Time 1 was after six months; time 2 after 18 months; and time 3 after 30 months.

[b] Autonomous group work at all three measurement occasions.

[c] Traditional job design at time 1; autonomous group work at times 2 and 3.

[d] Traditional job design at both measurement occasions.

[e] Traditional job design.

[f] Lower scores signify better mental health.

*$p<.01$

**$p<.001$

must reject the hypothesis of a work design effect on organizational commitment in this instance.

The final set of primary analyses concerned predicted changes over time within groups (Table 1 final row). Table 4 presents relevant findings. We obtained the predicted results for perceived autonomy of workgroups. Under condition B the mean score changed between the first and second measurement occasion and remained at the raised level at the final measurement. There was no corresponding change in either of the control conditions A and C.

The results for intrinsic job satisfaction show the equivalent pattern. In condition B, the recorded change was between the first and second measurement occasions, with the final mean remaining level. Again, no such change emerged for conditions A or C.

With respect to extrinsic job satisfaction, results are different. In condition B there was an increase in mean scores following the introduction of autonomous workgroups (time 1 to time 2) that diminished by the final

TABLE

Results of Analyses of Variance for Repeated

		Condition A[a]					
Dependent Variables	**Type of Analysis**	**Time 1[d]**	**Time 2**	**Time 3**	***F***	***df***	**Tukey Test**
Perceived work group autonomy	Repeated	45.9	49.4	49.7	1.0	2,20	–
	Independent	44.0	45.6	47.0	<1.0	2,124	–
Intrinsic job satisfaction	Repeated	37.6	36.4	34.9	1.2	2,36	–
	Independent	37.8	36.2	36.3	<1.0	2,145	–
Extrinsic job satisfaction	Repeated	47.3	43.8	41.4	6.5**	2,36	–
	Independent	46.8	44.0	43.5	3.3*	2,141	A1, A2 > A2, A3
Intrinsic job motivation	Repeated	36.1	36.5	33.4	5.6**	2,36	–
	Independent	35.2	34.0	34.3	<1.0	2,145	–
Organizational commitment	Repeated	52.4	52.1	45.0	6.3**	2,30	–
	Independent	50.7	48.1	46.3	3.0	2,138	A1, A2 > A2, A3
Mental health[e]	Repeated	18.6	20.8	23.7	5.8*	2,34	–
	Independent	18.8	19.6	21.0	3.2*	2,140	A2, A3 > A1, A2

[a] Autonomous group work at all three measurement occasions.

[b] Traditional job design at time 1; autonomous group work at times 2 and 3.

[c] Traditional job design at both measurement occasions.

[d] Means are reported. Time 1 was after six months; time 2 after 18 months; and time 3 after

[e] Lower scores signify better mental health.

*p<.05

**p<.01

measurement occasion. This increase reaches statistical significance for the analysis of variance for repeated measures but just fails to do so in the analysis of variance for independent groups ($F = 3.0, p < .06$). It is as if implementing this form of work design enhances extrinsic job satisfaction over the short rather than the long term. This interpretation is consistent with the findings for condition A, in which the initially high level of extrinsic satisfaction declined over time.

For the remaining three dependent variables, the third analysis again provided no evidence of an effect for autonomous workgroups. Results for condition B showed none of the predicted changes, and scores for the other conditions remained generally stable. This confirms both the first and second analyses. It is interesting to note that, in condition A, scores for mental health deteriorated over time. However, the effect is small, and the actual scores recorded at all three periods compare favorably with those of general populations (cf. Goldberg, 1978).

4

Measures and for Independent Groups over Time

Condition B^b						Condition C^c			
Time 1	Time 2	Time 3	*F*	*df*	Tukey Test	Time 1	Time 2	*F*	*df*
23.3	45.0	42.3	16.6**	2,16	–	31.4	30.8	<1.0	1,83
26.2	42.4	38.7	23.8**	2,84	B2, B3 > B1	31.6	31.8	<1.0	1,283
29.7	35.2	35.0	4.2**	2,16	–	30.0	28.9	2.3	1,113
31.6	34.6	35.4	4.1**	2,103	B2, B3 > B1	29.5	29.7	<1.0	1,344
39.4	43.8	40.0	3.6*	2,16	–	38.0	37.5	<1.0	1,114
41.1	44.7	42.3	3.0	2,103	–	37.5	38.1	<1.0	1,341
37.1	37.4	37.6	<1.0	2,16	–	31.5	32.3	<1.0	1,108
36.4	35.9	35.9	<1.0	2,107	–	31.3	32.3	1.8	1,348
48.3	52.0	51.2	1.6	2,18	–	41.6	41.7	<1.0	1,117
48.6	51.1	48.3	2.2	2,108	–	41.3	41.4	<1.0	1,354
21.1	19.9	23.9	1.5	2,14	–	21.7	21.4	<1.0	1,121
22.2	20.9	21.7	<1.0	2,103	–	22.1	21.4	2.4	1,353

30 months.

Interpreted with a strict line in causal inference, the results of this study provide evidence that the effects of autonomous workgroups are strong, but restricted to perceptions of workgroup autonomy and job satisfaction. Intrinsic job satisfaction shows an effect that persists over time, but it appears that the effect on extrinsic job satisfaction is less enduring. Autonomous workgroups have no clear effect on internal work motivation, organizational commitment, or mental health.

Supplementary Analyses

In a study such as this, random allocation across conditions cannot be achieved and labor turnover and other factors lead to missing values over time. Inevitably, rival explanations involving selection and mortality threaten causal interpretations. We consider selection issues first. Minor differences across groups existed from the outset in production processes and the ratio of men to women; in addition, there were differences in length of service. In order to examine the relevance of these threats, we repeated the three analyses described in the preceding sections with subsamples selected to be equivalent in these respects. Specifically, we conducted analyses in which (1) we restricted respondents from conditions C and D to those working on production processes that were virtually identical to processes performed by employees in conditions A and B; (2) we omitted the small number of men respondents; and (3) we omitted employees in conditions C and D whose length of service exceeded that of those in A and B. As in the three primary analyses, findings showed unambiguous effects for work design on perceived workgroup autonomy, intrinsic job satisfaction, and extrinsic job satisfaction, but no effects for intrinsic job motivation, organizational commitment, or mental health. Thus, differences between conditions in type of work, gender composition, and length of service cannot account for the findings.

Mortality is clearly a threat with respect to analysis of changes over time. The analysis of variance for repeated measures is necessarily based on a small proportion of the total population of respondents, and may therefore involve employees who do not represent the whole. Correspondingly, results from the analysis of variance for independent groups could provide spurious results because individuals lost or gained responded differently from those who remained throughout the study. To examine these threats we used t-tests to compare responses from newcomers and leavers in each experimental and control condition for each dependent variable with parallel responses for the same measurement occasions from employees also present at other time periods. None of these analyses yielded statistically significant differences. Thus we discounted threats based on mortality.

Performance, Labor Turnover, and Organizational Effects

From the outset an overall production target was set at the greenfield site. Based on criteria derived from work studies, this target specified the weight of output required according to the number of hours operators worked, called direct hours. In addition, during the first year the company set targets for individual and group output efficiency. Throughout this early period the amount produced by employees in conditions A and B fell far short of targets, individual and group efficiencies were low, and machine utilization was poor. With increased reliability of the technology, developing skill levels amongst the operators, and increased personnel levels, production steadily increased. Output rose from 50 tonnes to over 400 tonnes[3] per working month; overall efficiencies per working month exceeded 80 percent; quality improved and scrap decreased. Towards the end of the second year, employees in the two autonomous workgroup conditions reached output targets, and their outputs remained at that level for the remainder of the study and beyond. After three years the greenfield factory passed an assessment of its performance by the parent company.

However, direct comparison with the conventionally organized sister factory proved impossible. Throughout the study, output and efficiencies at this site remained stable and high. Data on production and performance were complex, and the criteria for the inclusion of different types of production information varied between sites. Some examples: standards for time required to turn out given products varied; each site categorized different employees as direct or indirect labor; and the established site only had information available at departmental or factory levels, not for individuals.

Nevertheless, examination of company records and discussions with relevant personnel did provide some insights into the meanings of the production data. First, these people felt that overall efficiency at the greenfield site was approximately 5 to 10 percent lower than that at the comparison site. However, they did not see lack of effort from individual operators as the major cause; rather, they saw employee performance at the two sites as broadly comparable. Moreover, at the greenfield factory, the target system largely determined output levels and employees' efforts, and there was no requirement that operators should exceed the targets. Second, in attempting to maintain efficiency levels, the factories responded differently to market demands. Managers at the conventional site recruited or released short-term labor when demand fluctuated. The greenfield site maintained stable personnel levels as far as possible in order to protect and foster the spirit and practice of team working. By the end of the study the autonomous workgroups could easily meet increased market demands by raising ouput and efficiency. Thus, the two sites implemented different organizational strategies to maintain the efficient production of sweets. Neither site could be proved to be superior in performance. This was contrary to predictions for individuals and groups at the greenfield site.

One implication of this discussion of productivity is that productivity benefits could accrue at the greenfield site once targets had been met. This is because there were no indirect costs for supervisory staff. The established site had on the average one supervisor for every ten shopfloor operators, one of whom acted as chargehand in the supervisor's absence. Productivity gains at the greenfield site could of course be lost during periods of low demand when direct labor was underused.

---

[3] A tonne = 1000 kg. (approximately 2240 lbs.) in the metric system of weights.

Contrary to predictions, examination of the data on labor turnover showed that the overall incidence for the autonomous work conditions (28% annually) exceeded that for the comparison conditions (22%). Analysis of the proportion of leavers to stayers over two-and-a-half years of the investigation yielded a statistically significant difference ($\chi^2 = 52.7, df = 1, p < .001$). Breaking this overall figure down according to reasons for leaving produced an interesting picture. With respect to resignations for personal/domestic and health reasons, there was no difference between autonomous and traditional work design conditions ($\chi^2 = 1.75$, n. s.; $\chi^2 < 1$, n. s.). By comparison, the proportion of those who left for other jobs was higher in the autonomous work conditions ($\chi^2 = 19.72, df = 1, p < .001$). It seems unlikely this indicates general dissatisfaction as analyses showed no difference between leavers and stayers on job satisfaction or other attitudinal measures. It is more likely that this finding reflects differences in local labor markets because the autonomous workgroups were in a region having lower levels of unemployment than the conventional site's region. In addition, some individuals may not fit into a system of group working. But the most marked difference between the conditions with autonomous and conventional job designs concerned dismissals for disciplinary reasons. Labor turnover in this category was at a rate of 8.2 percent in the autonomous workgroups, whereas in the comparison groups it was only 1.3 percent. The difference between the two was highly significant statistically ($\chi^2 = 135.25, df = 1, p < .001$), though the absolute number of dismissals was small.

One explanation for this finding is that rate of dismissals will go up following any major recruitment program since a small proportion of employees will not fit easily into their new organization. However, there is some evidence that the high number of disciplinary dismissals was also a direct consequence of autonomous group working. Interviews with managers and employees support this interpretation and illustrate some of the organizational implications of sustaining autonomous workgroups. Line managers were aware that they had dismissed employees who, under conventional circumstances, they would not have required to leave. They attributed their taking a harder line in part to the pressures on them to produce and to develop the new site, but they saw a side effect of autonomous group working as the main reason. The absence of supervisors meant that the few employees who stepped out of line or abused their freedom were not immediately detected. Under conventional job designs, supervisors would take early action and perhaps protect employees from managers. In autonomous workgroups, misdemeanors were typically more advanced when they came to light and warranted use of formal warning procedures. In other words, management at the greenfield site relied more on formal procedures and sanctions than did the conventional site's managers. Interestingly, the workgroups were loath to discipline members who consistently broke group norms and explicitly referred these problems to the support managers.

Some findings from the interview program conclude this section. At the start-up of production at the greenfield site, managers reported high levels of stress, and interpersonal and interfunctional conflict. These reactions were due to the difficulties involved in developing a greenfield site and managing an untried system of group working. Indeed, some managers believed that the experiment with group working made it a harder and slower job to get the new site working effectively. Over time, with the solving of technical problems and the emerging success of the work groups, the managers' stress levels fell, their attitudes toward their jobs improved,[4] and the management team became more cohesive. However, they still perceived their roles as highly stressful. As one manager stated, "The mental, physical, and emotional pressure on the factory management with this type of organization should not be underestimated and, dare I suggest, is significantly greater than for our equivalents in other factories." Indeed, the support managers felt the greatest level of pressure both in terms of the production output required and in terms of the difficulties involved in communicating, intervening, disciplining, and providing the resources and conditions necessary for up to 70 employees in six teams to work together effectively and efficiently. As a result the factory employed four different day-shift support managers in the first three years; this has now stabilized. The difficulty was exacerbated because it was not possible to promote personnel within the factory via supervisory positions. Hence, new job incumbents needed time to be socialized into the organization but immediately experienced the problems of managing semiautonomous workgroups.

But in pointing out these problems, we should not lose sight of the benefits of autonomous workgroups. In this study, these benefits were sufficient to ensure that operators and managers alike supported the system, which is to be continued for the foreseeable future. During interviews, for example, we saw clearly that shopfloor employees actively enjoyed the system and its attendant responsibilities. None preferred conventional forms of work design of which all but a few had prior experience. All were pleased not to have supervisors breathing down their necks and spoke very highly of their work conditions. Employees at the new factory particularly enjoyed not having to clock on and off and also appreciated the various other ways in which management treated them as mature adults. As one employee stated, "After 20 years of working at two other firms, the conditions here are excellent—managers treat you as people." Indeed, another suggested that "work groups enable us to get things off our chests and do something about it." One consequence of this satisfaction was a good industrial relations climate in the factory.

---

[4] Managers completed the same work attitude measures and General Health Questionnaire administered to the shopfloor employees (described in the Method section) both at the start-up of production and 30 months later.

DISCUSSION AND CONCLUSIONS

This field experiment showed that autonomous workgroups had specific rather than wide-ranging effects on employee attitudes and behavior. Although implementation of this work design substantially enhanced intrinsic job satisfaction, it did not demonstrably affect job motivation, organizational commitment, mental health, work performance, and voluntary labor turnover. At an organizational level, implementation made productivity benefits possible because the autonomous workgroups reduced indirect labor costs; a side effect was increased dismissals. Employees clearly appreciated the autonomous work system. On balance, managers did too, though clearly there were costs in terms of personal stress arising from the difficulties involved in managing and maintaining the system.

Overall, the most striking aspect of these findings is the restricted range of effects on individual attitudes compared with the range predicted from the literature. One possible explanation for this is that the form of work design introduced did not differ sufficiently from conventional practice. However, this is implausible. Objectively, the system of autonomous workgroups was a radical departure for this organization, and the results pertaining to employees' perceptions of autonomy of workgroups testify that those involved experienced it as such. As in any investigation, a second possible explanation is that the findings are idiosyncratic. This possiblity can never be logically dismissed but seems implausible. Work design theory applies generally to shopfloor jobs, and there is no reason to dismiss the present setting as atypical in this regard. The most likely explanation for the restricted range of effects on employee attitudes is that the rigorous research design eliminated causal interpretations that weaker designs would have supported. Even within the present study it is easy to see how this could have occurred with a different research design. If for example we had restricted the design to conditions A and C, we would have inferred effects on job motivation, organizational commitment, and mental health, as Table 4 makes evident. Much of the previous research from which we took predictions is of such a comparative nature. The general point is that more rigorously designed studies serve to eliminate a greater number of hypotheses.

An interesting aspect of the present study is that it analyzed two different approaches to implementation of autonomous group working. Employees in condition A were hired for jobs under the new form of work design, whereas employees in condition B had their jobs redesigned. As Roberts and Glick commented, "the act of having one's job changed may have different influences on a person than does his or her signing on to impoverished or complex tasks" (1981: 210). However, the results of the present study show quite clearly that both approaches had the same effect. Both improved job satisfaction. This suggests that it is the nature of autonomous workgroups that is the main causal factor in determining employees' attitudes.

The finding that the satisfaction effect is stronger and more enduring for intrinsic rather than extrinsic aspects is interesting. This is consistent with theory that emphasizes the importance of job content to people's experience of work itself. However, the finding also suggests that satisfaction ceases to overflow to wider contextual circumstances over time. This reflects earlier theory (Herzberg, 1966) ascribing secondary importance to conditions of employment.

The most problematic outcome of this study is the absence of effects for intrinsic job motivation. Implicitly or explicitly, this concept is at the core of most approaches to job design (Davis, 1966; Hackman & Oldham, 1976).[5] In the present case it might be argued that failure to find such an effect resulted from the use of an insensitive instrument. The intrinsic job motivation measure yielded responses distributed towards the top of the scale, indicating that a ceiling effect was to some extent operative. Although this distribution supports the insensitivity argument, it is not compelling, since the scale was sensitive enough to reveal statistically significant differences across conditions and over time that did not correspond to effects of work design. On the basis of this study, it would seem that no demonstrable motivational benefit derives from autonomous group working. To the extent that autonomous workgroups reflect some general principles of job design this finding is of considerable significance in suggesting that major theoretical tenets may be invalid.

The absence of an unambiguous effect on organizational commitment is less troublesome given recent theoretical developments. Though earlier work (Emery, 1959) led us to expect such an effect, which was also an outcome the company desired, recent theory (Hackman, 1983) suggested that autonomous workgroups may have their effect on commitment to workgroups rather than to organizations. Our investigation did not address this possibility, which deserves further empirical attention. Similarly, job design had no effect on the other exploratory variable, mental health. It may be, however, that this variable is still important in the context of impoverished jobs (Broadbent, 1985).

Current theory also offers several reasons for the lack of a performance effect. The factory we studied did not implement all the design criteria for autonomous workgroups that current theory has made explicit (e. g. Hackman, 1983). In particular, rewards were distributed to individuals rather than groups, performance feedback tended to mix information between the individual and group levels, and resources for training were low. It is possible that these deficiencies prevented the autonomous workgroup system from having its predicted direct effect on performance.

We turn now to the practical implications. The results show that autonomous group working is a viable proposition that is appreciated by those who experience it. Moreover, it may have clear economic benefits. This is not, as theory predicts, because groups enhance operators' motivation and effort; economic benefits stem instead from the logic of the groups themselves. With decision making delegated to shopfloors, the need for supervision declines or vanishes, indirect labor costs decrease, and productivity benefits can accrue. Given appropriate organizational designs, resources, and managerial structures, task performance might also be improved (Hackman, 1983). But it is also clear that such work systems have their costs, albeit ones that are difficult to quantify because they chiefly involve managerial time and effort. Furthermore, even these apparently straightforward practical implications need to be qualified. It should not be forgotten that the findings obtained were from a greenfield setting. This setting had the advantage of allowing the implementation of a radical and well-developed form of autonomous workgroups. But by the same token this setting limits generalizability. It is reasonable to assume that in established organizations the threat autonomous workgroups pose for the nature and continued existence of supervisory jobs might reduce or outweigh their potential practical benefits. Similarly, the emphasis on employees having many skills may cut across agreed demarcation lines and customs and practices of work, particularly in unionized settings. In established settings, people might resist such change. Thus, in such contexts, the process of change may be as important as its content.

---

[5] See Clegg (1984) for an alternative theoretical account.

REFERENCES

Alcalay, R., & Pasick, R. J. 1983. Psycho-social factors and the technology of work. ***Social Science and Medicine***, 17: 1075–1084.

Banks, M. H., & Jackson, P. R. 1982. Unemployment and risk of minor psychiatric disorder in young people: Cross-sectional and longitudinal evidence. ***Psychological Medicine***, 12: 789–798.

Banks, M. H., Clegg, C. W., Jackson, P. R., Kemp, N. J., Stafford, E. M., & Wall, T. D. 1980. The use of the General Health Questionnaire as an indicator of mental health in occupational studies. ***Journal of Occupational Psychology***, 53: 187–194.

Broadbent, D. E. 1985. The clinical impact of job design. ***British Journal of Clinical Psychology***, 24: 33–44.

Campbell, D. T., & Stanley, J. C. 1966. ***Experimental and quasi-experimental designs for research***. Chicago: Rand McNally & Co.

Clegg, C. W. 1984. The derivation of job designs. ***Journal of Occupational Behaviour***, 5: 131–146.

Clegg, C. W., & Wall, T. D. 1981. A note on some new scales for measuring psychological well-being at work. ***Journal of Occupational Psychology***, 54: 221–225.

Cook, J. D., & Wall, T. D. 1980. New work attitude measures of trust, organizational commitment and personal need non-fulfilment. ***Journal of Occupational Psychology***, 53: 39–52.

Cook, J. D., Hepworth, S. J., Wall, T. D., & Warr, P. B. 1981. ***The experience of work: A compendium and review of 249 measures and their use***. London: Academic Press.

Cook, T. D., & Campbell, D. T. 1979a. ***Quasi-experimentation***. Chicago: Rand McNally & Co.

Cook, T. D., & Campbell, D. T. 1979b. Four kinds of validity. In R. T. Mowday & R. M. Steers (Eds.), ***Research in organizations***: 77–101. Santa Monica, Calif.: Goodyear.

Cummings, T. G. 1978. Self-regulating work groups: A socio-technical synthesis. ***Academy of Management Review***, 3: 625–634.

Cummings, T. G., Molloy, E. S., & Glenn, R. 1977. A methodological critique of fifty-eight selected work experiments. ***Human Relations***, 30: 675–708.

Davis, L. E. 1966. The design of jobs. ***Industrial Relations***, 6: 21–45.

Davis, L. E., & Valfer, E. S. 1965. Intervening responses to changes in supervisors' job designs. ***Occupational Psychology***, 39: 171–189.

Denison, D. R. 1982. Sociotechnical design and self-managing work groups: The impact of control. ***Journal of Occupational Behaviour***, 3: 297–314.

Emery, F. E. 1959. ***Characteristics of sociotechnical systems***. London: Tavistock Institute of Human Relations.

Emery, F. E. 1980. Designing socio-technical systems in "greenfield" sites. ***Journal of Occupational Behaviour***, 1: 19–27.

Fraser, R. 1947. ***The incidence of neurosis among factory workers***. Report No. 90, Industrial Health Research Board, HMSO, London.

Gardell, B. L., 1971. Alienation and mental health in the modern industrial environment. In L. Levi (Ed.), ***Society, stress and disease***, vol. 1: 148–180. New York: Oxford University Press.

Goldberg, D. P. 1972. ***The detection of psychiatric illness by questionnaire***. Oxford: Oxford University Press.

Goldberg, D. P. 1978. ***Manual for the General Health Questionnaire***. Windsor, U. K.: National Foundation for Educational Research.

Gulowsen, J. 1972. A measure of work group autonomy. In L. E. Davis & J. C. Taylor (Eds.), ***Design of jobs***: 374–390. Harmondsworth, U. K.: Penguin.

Hackman, J. R. 1977. Work redesign. In J. R. Hackman & J. L. Suttle (Eds.), ***Improving life at work***: 93–162. Santa Monica, Calif.: Goodyear.

Hackman, J. R. 1983. The design of work teams. In J. Lorsch (Ed.), ***Handbook of organizational behavior***: 70–94. Englewood Cliffs, N. J.: Prentice-Hall.

Hackman, J. R., & Lawler, E. E. 1971. Employee reactions to job characteristics. ***Journal of Applied Psychology***, 55: 259–286.

Hackman, J. R., & Oldham, G. R. 1976. Motivation through the design of work: Test of a theory. ***Organizational Behavior and Human Performance***, 15: 250–279.

Hannan, M. T., & Young, A. A. 1977. Estimation in panel models: Results on pooling cross-sections and time series. In D. Heise (Ed.), ***Sociological methodology***: 52–83. San Francisco: Jossey-Bass.

Herbst, P. G. 1974. ***Socio-technical design***. London: Tavistock Publications.

Herzberg, F. 1966. ***Work and the nature of man***. New York: World of Publishing Co.

Jackson, P. R., Stafford, E. M., Banks, M. H., & Warr, P. B. 1983. Unemployment and psychological distress in young people: The moderating role of employment commitment. ***Journal of Applied Psychology***, 68: 525–535.

Karasek, R. A. 1978. Job demands, job decision latitude, and mental strain: Implications for job design. ***Administrative Science Quarterly***, 24: 285–308.

Katz, D., & Kahn, R. L. 1978. ***The social psychology of organizations*** (2nd ed.). New York: John Wiley & Sons.

Kelly, J. E. 1978. A re-appraisal of sociotechnical systems theory. ***Human Relations***, 31: 1069–1099.

Kemp, N. J., Clegg, C. W., & Wall, T. D. 1980. Job redesign: Content, process, and outcomes. ***Employee Relations***, 2(5): 5–14.

Kemp, N. J., Wall, T. D., Clegg, C. W., & Cordery, J. L. 1983. Autonomous work groups in a greenfield site: A comparative study. ***Journal of Occupational Psychology***, 56: 271–288.

Kornhauser, A. W. 1965. ***Mental health of the industrial worker.*** New York: John Wiley & Sons.

Lawler, E. E. 1977. Adaptive experiments: An approach to organizational behavior research. ***Academy of Management Review,*** 2: 576–585.

Nerlove, M. 1971. Further evidence on the estimation of dynamic economic relations from a time series of cross-sections. ***Econometrics,*** 39: 359–382.

Pasmore, W., Francis, C., Haldeman, J., & Shani, A. 1983. Sociotechnical systems: A North American reflection on the empirical studies of the seventies. ***Human Relations,*** 35: 1179–1204.

Rice, A. 1958. ***Productivity and social organization: The Ahmedabad experiment.*** London: Tavistock Publications.

Roberts, K. H., & Glick, W. 1981. The job characteristics approach to job design: A critical review. ***Journal of Applied Psychology,*** 66: 193–217.

Rousseau, D. M. 1977. Technological differences in job characteristics, employee satisfaction and motivation: A synthesis of job design research and socio-technical systems theory. ***Organizational Behavior and Human Performance,*** 19: 18–42.

Simonton, D. K. 1977. Cross-sectional time-series experiments: Some suggested statistical analyses. ***Psychological Bulletin,*** 84: 489–502.

Susman, G. I. 1976. ***Autonomy at work.*** New York: Praeger.

Trist, E. L., & Bamforth, K. 1951. Some social and psychological consequences of the long wall method of coal-getting. ***Human Relations,*** 1: 3–38.

Trist, E. L., Susman, G., & Brown, G. W. 1977. An experiment in autonomous group working in an American underground coal mine. ***Human Relations,*** 30: 201–236.

Wall, T. D., & Clegg, C. W. 1981. A longitudinal field study of group work redesign. ***Journal of Occupational Behaviour,*** 2: 31–49.

Wall, T. D., Clegg, C. W., & Jackson, P. R. 1978. An evaluation of the Job Characteristics Model. ***Journal of Occupational Psychology,*** 51: 183–196.

Warr, P. B., Cook, J. D., & Wall, T. D. 1979. Scales for the measurement of some work attitudes and aspects of psychological well-being. ***Journal of Occupational Psychology,*** 52: 129–148.

Winer, B. J. 1971. ***Statistical principles in experimental design*** (2nd ed.). New York: McGraw-Hill Book Co.

APPENDIX A

The questionnaire used in this study included the following four scales to measure attitudes towards work (Warr, Cook, and Wall, 1979). All items had appropriately worded 7-point response formats.

The intrinsic job satisfaction scale had seven items. Examples are:
How satisfied are you with—
the freedom to choose your own method of working?
the amount of responsibility you are given?
the opportunity to use your abilities?

The extrinsic job satisfaction scale had eight items. Examples are:
How satisfied are you with—
the physical work conditions?
your rate of pay?
the way your firm is managed?

The intrinsic job motivation scale had six items. Examples are:
Do you agree or disagree with the following—

I take pride in doing my job as well as I can.
My opinion of myself goes down when I do this job badly.
I try to think of ways of doing my job effectively.

The organizational commitment scale had nine items. Examples are:
Do you agree or disagree with the following—
I am quite proud to be able to tell people who it is I work for.
I feel myself to be part of the organization.
In my work I like to feel I am making some effort, not just for myself but for the organization as well.

APPENDIX B
Statistical Analyses

Generalized least squares. The design of this study had three main features that precluded the use of standard methods for analysis of variance. First, the repeated measures element of the panel design had to be incorporated in order to (1) increase the power of statistical tests, and (2) provide for the repeated measures that, if ignored, would have biased significance tests. Second, the growth in personnel levels at the greenfield site together with a reduction in the workforce at the established site necessarily resulted in a large number of missing values (38%). Finally, the study did not constitute a full factorial design. Autonomous workgroups were introduced in one site only, leading to partial confounding between site and work design factors.

What was required was an approach that (1) capitalized on the information inherent in the repeated measures element, (2) used the independent groups component appropriately, and (3) coped with the design's lack of balance. The method of analysis chosen was based on the so-called cross-sectional time series designs (Hannan & Young, 1977; Nerlove, 1971) discussed in the econometric literature. Essentially, they give a split-plot type of ANOVA design, with a between-group term containing the hypotheses of interest and a within-group term that provides for repeated measurement.

Hannan and Young (1977) and Simonton (1977) discussed several approaches to parameter estimation, but the most common method is to use generalized least squares after estimating the autocorrelation parameter from the data.

The first step of estimating the autocorrelation parameter was applied only to data for those respondents present on all three measurement occasions, with respondents represented by dummy variables (Hannan & Young, 1977). We used the regression weights for the respondents' dummy variables to calculate the autocorrelation coefficient by the formula given by Hannan and Young (1977: 65). We then adjusted the raw data from the whole set of respondents on the basis of the derived autocorrelation estimate. Finally, we ran ordinary-least-squares regressions on the adjusted scores for the total sample to give generalized-least-squares estimates that are unbiased estimates of the terms of interest and their standard errors. At this final stage, we used dummy variables to code site, shift, and type of work design. Measurement period was also included in the equation.

One-way analysis of variance. The generalized-least-squares analysis has the advantage of parsimony; for each dependent variable it provides an omnibus test of relevant hypotheses using the whole data set with statistical control of relevant parameters. However, it does not provide directly for the clear examination of the specific patterns of predictions across conditions or over time. The second stage in the analysis consequently involved analysis of variance applied in two steps.

First, one-way analysis of variance for independent groups was used for each dependent variable at each measurement occasion separately. Where a statistically significant effect was obtained, we applied an a posteriori Tukey contrast test (Winer, 1971) with a prescribed significance level of $p < .01$.

The second and final step in the analysis concerned change over time within each condition. We conducted both analysis of variance for repeated measures (for those respondents with valid data at all measurement occasions) and analysis of variance for independent groups for each dependent variable. The repeated-measures analysis is potentially the more sensitive of the two, but has far fewer degrees of freedom because of missing values; the independent-groups analysis is more conservative since the error term is inflated by the inclusion of subject variance, but at the same time encompasses all respondents. We applied a posteriori contrast tests as before.

Together, these two analyses provided a useful way of handling the data, subject to checks on the representativeness of the repeated-measure subsample and of distortions due to mortality effects in the independent-groups analysis. The Results section describes the outcome of these checks.

Toby Wall obtained his Ph. D. degree from Nottingham University. He is currently Assistant Director at the Medical Research Council (MRC) and Economic and Social Research Council (ESRC) Social and Applied Psychology Unit, at University of Sheffield, England. His research interests cover information technology in medical and manufacturing organizations; work design; employee motivation, mental health, and well-being; and applied research design.

Nigel Kemp was awarded his Ph. D. degree by the University of Kent at Canterbury. He is a lecturer in management sciences at the University of St. Andrews, Scotland. His current research interests include the psychological and organizational effects of information technology, work design and QWL, nonverbal communication, and the social psychology of blindness.

Paul Jackson received his Ph. D. degree from the University of Sheffield and is currently a research fellow at the MRC/ESRC Social and Applied Psychology Unit, University of Sheffield, England. His research interests are unemployment, social support systems, and statistical methods.

Chris Clegg was awarded his M. A. degree by the University of Bradford. He is a senior research fellow at the MRC/ESRC Social and Applied Psychology Unit, University of Sheffield, England. His research interests are the human aspects of information technology, absence and labor turnover, and the impact of uncertainty in organizations.

© *Academy of Management Journal*
1986, Vol. 29, No. 2, 305–328.

EFFECTS OF TEAM BUILDING AND GOAL SETTING ON PRODUCTIVITY: A FIELD EXPERIMENT

PAUL F. BULLER
Utah State University
CECIL H. BELL, JR.
University of Washington

A field experiment conducted with 53 hard-rock miners in an underground metal mine examined the effects of two interventions, team building and goal setting, on miners' productivity and strategy development. The study employed a 2 × 2 factorial design with random assignment of participants to the goal setting condition, and it used objective criteria of performance. Unfortunately, the results were inconclusive. Although slight improvements occurred in some measures of performance and strategy development, it is not clear that these improvements were results of the interventions. More of these improvements occurred in the goal setting than in the team building condition. Limitations in the research design and factors in the experimental setting that could not be controlled led to equivocality of these results.

Numerous techniques drawn from behavioral science for improving individual and organizational performance figure in the contemporary management scene. Practicing managers are offered a veritable smorgasbord ranging from alphabet soup (QWL; OD; MBO; TA; OB Mod; Theories X, Y, Z; QC; PDM; etc.) to nuts. Naturally, their adherents and practitioners make claims for the efficacy of these techniques and counterclaims are forthcoming from critics and skeptics. Three reviews of the empirical literature (Cummings, Molloy, & Glen, 1975; Huse, 1980; Locke, Feren, McCaleb, Shaw, & Denny, 1980) generally agreed that only a few intervention techniques have been shown to be effective through well-designed and well-executed research investigations. We believe that rigorous evaluation of applied behavioral science programs is desirable and that field experiments in organizational contexts are particularly needed. This study describes a field experiment that examined the effects of two prominent intervention techniques, team building and goal setting, on individuals' productivity.

TEAM BUILDING

One of the most popular intervention techniques in organization development (OD) is team building (French & Bell, 1984; Porras & Berg, 1978). The

---

The authors wish to thank Gary P. Latham for his helpful comments on an earlier version of this paper. The suggestions of three anonymous reviewers were also most helpful. This research was conducted in part under U.S. Bureau of Mines contract number J0387230.

primary objective of team-building activities is to increase the effectiveness of work teams. This is usually achieved through improving problem solving and interpersonal relations in teams. Team building is not a well-defined concept, however. Organizations employ many variations; the emphases of programs change over time; and programs are often situation- and problem-specific. Beer (1976) described four models of team building: a goal-setting model, an interpersonal relations model, a role clarification model, and the managerial grid model. Dyer (1977) described a more general model of team building based on problem solving in which team members identify major problems, generate relevant information, engage in problem solving and action planning, and implement and evaluate action plans. We refer to this task-oriented problem-solving technique for teams as team building/problem solving. The assumption underlying this type of intervention is that work teams become more effective by solving their major problems together. The team-building program studied in this research used team building/problem solving techniques that are operationally defined in the Methods section.

Research assessing the effectiveness of team building has been inconclusive (DeMeuse & Liebowitz, 1981; Nicholas, 1982; Woodman & Sherwood, 1980). Positive results were found in 29 of the 36 studies (80%) reviewed by DeMeuse and Liebowitz, in 2 of the 4 studies reviewed by Nicholas, and in 19 of the 30 studies (63%) reviewed by Woodman and Sherwood. Nevertheless, these authors concluded that they could make no firm statements regarding the efficacy of team building with respect to its effects on the performance of individuals and work groups. They based their conclusions on the fact that the research designs of the studies reviewed were generally weak, and the findings thus lacked internal validity. Moreover, with the exception of some reported by Nicholas, the studies measured subjective data such as employees' perceptions and attitudes and usually ignored objective data on performance. Finally, they often vaguely and imprecisely described interventions.

Buller's (1982) review of the literature on team building yielded only nine studies that implemented clearly defined team-building interventions and measured effects on objective criteria of performance. Of these nine studies, six reported positive effects on performance. However, all but one (Margulies & Raia, 1968) of the successful team-building interventions were confounded with goal setting or other activities; thus it was impossible to determine which technique produced the results. Margulies and Raia's study used a preexperimental design and thus had limited internal validity.

An additional problem with previous research on team building is its failure to specify and test relationships between independent variables, causal mechanisms, and criteria of performance. It is simply not clear why team building affects performance, if it does at all. Nicholas (1982) suggested that team building may lead to improved performance through "modification of attitudes, values, problem-solving techniques, and interpersonal and group processes" (1982: 535). A recent model presented by Buller (1983) suggested that team building potentially affects a number of individual, group, task, and organizational factors that in turn improve performance. This model predicts that the performance of a team will improve to the extent that it can identify and solve major debilitating problems in one or more of these areas. However, existing models are weak in defining specific theoretical relations. Perhaps Hackman's (1976) model of the effects of workgroups on the performance of individuals is the most relevant to this study and the most parsimonious in describing how team building might work. He argued that individual performance is a function of an individual's ability, motivation, effort, and strategies for performance—all moderated by the nature of the task being performed. To the extent that immediate workgroups enhance or hinder these mechanisms, they enhance or hinder performance. Certain task contingencies may affect which of the causal mechanisms is most important to improved performance.

This research was not an attempt to develop or validate a theory of team building. The primary purpose of the study was to test team building's independent effects on performance using objective criteria of performance. However, the research directly examined the effects of team building on development of task performance strategies, an individual cognitive mechanism. To our knowledge, no previous research has directly tested such a cognitive effect of team building.

GOAL SETTING

The importance of goal setting for motivating employees' performance was first explicitly recognized in the management literature by Taylor (1911); Locke (1968) translated the concept into a contemporary theory of task performance. According to Locke's theory, if workers accept hard, specific goals, performance improves. In the most thorough review of goal setting to date, Locke, Shaw, Saari, and Latham (1981) viewed goal setting as primarily a motivational phenomenon, but acknowledged that it also involves cognitive elements. From a motivational standpoint, goal setting affects the direction, amplitude (effort), and duration (persistence) of action. From a cognitive perspective, goal setting is thought to affect strategy development or ways to accomplish goals. The first three of these mechanisms are relatively direct in their effects on performance. Strategy development is more indirect. Locke and colleagues (1981) noted that this concept

> involves developing strategies or action plans for attaining one's goals. Although strategy development is motivated by goals, the mechanism itself is cognitive in essence; it involves skill development or creative problem solving (1981: 132).

Goal setting, the subject of extensive experimental research in both laboratories and field settings, has been found to be a robust technique for improving task performance (Latham & Yukl, 1975; Locke et al., 1981). Most of this research has focused on the first three mechanisms by which goals affect performance. Few studies have directly examined the effects of goal setting on strategy development, although several studies have inferred such effects. For example, Bandura and Simon (1977) noted that overweight clients in a weight clinic who had specific goals for mouthfuls eaten altered their eating habits. Some individuals ate low calorie foods that did not count in their quotas; some engaged in more planning than before. Latham and Baldes (1975) observed similar behavior in a study of drivers of logging trucks. Some of the drivers who were assigned specific challenging goals for trucks' weights recommended small changes in their trucks that let them more accurately judge weight. In general, participants in these studies with hard specific goals engaged in more creative problem solving than participants with no goals.

In another relevant study, Terborg (1976) obtained separate measures of the effects of goal setting on individuals' efforts and the direction of their behavior on a reading task in a simulated organization. The study operationally defined direction of behavior as the extent to which individuals employed appropriate behaviors or action plans to accomplish the task. This was essentially a measure of strategy development. The results showed that goal setting was related to both effort and direction of behavior, which in turn were related to performance. Specifically, individuals who set hard specific goals on the reading task spent more time on the reading materials (effort) and were more likely to use learning strategies relevant to their goals (direction of behavior) than those who did not set goals.

Finally, scholars have observed that strategy development may be especially important in complex tasks because if people do not develop strategies that are relevant to goals, increased motivation caused by goal setting may not enhance performance (Bavelas & Lee, 1978; Christensen-Szalanski, 1980). The nature of tasks may therefore be an important variable determining whether or not goal setting will stimulate development of strategies, lead to improved performance, or both.

The studies discussed in the preceding paragraphs support conceptual arguments that both cognitive and motivational mechanisms influence performance of tasks (Naylor, Pritchard, & Ilgen, 1980; Terborg & Miller, 1978). They provide some evidence that goal setting affects cognitive as well as motivational determinants of performance. To date, only one reported study has systematically examined the effects of goal setting on strategy development. In a laboratory experiment, Shaw (1983) found that specific hard goals led to the development of more task strategies than no goals did. The present study investigated the effects of goal setting on strategy development and task performance in a field setting.

In summary, our hypotheses were straightforward: team building and goal setting will each have significant, positive effects on task performance and strategy development.

METHODS

Context and Setting

The present research was part of a larger organization development (OD) project designed to demonstrate the effectiveness of OD techniques for improving safety and productivity in mines. Conducted at a silver mining company, the program consisted primarily of team-building activities at all levels of the organization but also involved a variety of other activities. These included a revision of the company's performance appraisal system to place more emphasis on safety, an assessment of data processing needs, development of a program for training supervisors, and the experiment reported in this paper. Although the other activities potentially confounded the effects of the experiment, it is important to note that only the supervisory training program also directly involved hourly workers. Further, the organization did not implement this training program until after we had completed our study. The entire OD project took approximately three years; the field experiment extended over eight months.

The site for the field experiment was an underground mine producing silver, lead, and zinc. The mine employed 292 persons with approximately 220 hourly employees involved in underground production operations. The average hourly employee was 38 years old, had a twelfth grade education, and had 12 years of work experience. About 60 percent of the workforce belonged to an international union. The mine was a hard-rock mine in which ore-bearing rock in vein structures was mined in small working places called stopes.

As part of the overall OD program, a team-building intervention was conducted with approximately half of the hourly underground employees and their supervisors at the mine. Adding the goal-setting induction generated a 2 × 2 factorial design and thereby allowed the controlled investigation of both behavioral science interventions in a single field experiment. Thus, the experiment was a study within a study and was more limited in scope than the overall OD effort. To our knowledge no one has previously implemented and evaluated either team building or goal setting in the mining industry. Additionally, few studies of team building have been conducted with hourly employees as participants.

Study Design

The units of analysis for this study were stopes, small underground working places in the mine. Twenty of the mine's 23 production stopes were assigned to the treatment conditions; 3 stopes were excluded because they were not operational during the course of the experiment. To implement the team-building condition, the mine superintendent nonrandomly assigned 12 stopes consisting of 36 miners to undergo team building; the 8 remaining stopes were assigned to the no-team-building condition. The reasons the 12 stopes were assigned to team building were: (1) four intact work crews operated them, each crew consisting of a shift boss (first-line supervisor), stope miners, and hourly support personnel; (2) they constituted two complete production levels underground; (3) they included opposite shift crews, with one crew working the day shift and the other crew working the afternoon shift on the same level, and the superintendent wanted to address problems that existed between opposite shifts; (4) the locations of these stopes were important to overall mine production; and (5) the superintendent wanted at least half of the hourly workforce to get an exposure to team building.

Two weeks after the team-building intervention was completed, the goal-setting condition was implemented by randomly assigning 10 stopes consisting of 26 miners to the goal-setting condition and 10 stopes consisting of 27 miners to the no-goal-setting condition. Thus, a total of 53 miners, all men, in 20 production stopes were assigned to experimental conditions as shown in Figure 1. The resulting 2 × 2 factorial design was quasi-experimental because it lacked complete random assignment to conditions and simultaneous induction of the two treatments (Cook & Campbell, 1979). Figure 2 shows the timeline for the experiment.

Two miners worked most stopes, but some stopes had one, three, or four individuals. Since it was operationally impossible to determine every miner's individual performance, we used performance of stope teams as the unit of analysis. Clearly, this situation presents problems in distinguishing between the performance of individuals and groups.

The stope miners' task. Stope miners are the primary production workers and major source of increases or decreases in productivity in such mines. Typically, two miners work in each stope for long periods of time, working both alone and together. Each mines a horizontal cut about 100 feet long in opposite directions from the raise, a vertical opening on the vein. All other times they help each other in certain phases of the mining cycle. Miners share equal responsibility for a stope's production. As the rock is mined out, they pump a sand slurry back in to raise the floor of the stope and the mining cycle of "drill, blast, and muck"[1] begins again. All miners use similar equip-

FIGURE 1
Factorial Design for the Experiment

	Goal Setting	No Goal Setting
Team Building	n = 6 stopes (18 miners)	n = 6 stopes (18 miners)
No Team Building	n = 4 stopes (8 miners)	n = 4 stopes (9 miners)

ment and mining methods. A crew of support people assists the stope miners by bringing necessary supplies to the work area and helping to remove the ore. Each level of the mine has from one to six working stopes, and one support crew serves all stopes on a level.

Contact between Researchers and Participants

Sponsored by the U.S. Bureau of Mines as a consultant for the three-year OD project, the second author conducted all of the team-building sessions described later in this section. The first author, also sponsored by the Bureau of Mines, was a research assistant during one year of the OD project. As a research assistant, he observed but did not participate in the team-building sessions with underground crews, conducted the goal-setting intervention, and collected and analyzed the data reported here.

Independent Variables

The experimental treatments were team building/problem solving and goal setting. Each variable had two conditions, the presence or absence of the treatment.

Team building/problem solving. This variable was operationally defined as a planned series of meetings, facilitated by an outside consultant, with a team—a group of people having common organizational relationships and goals. These meetings must be designed specifically to improve the team's accomplishment of tasks by (1) developing problem-solving procedures and skills and (2) then solving the team's major problems.

The team-building induction was implemented through a series of six 45-minute meetings over a period of ten weeks. Entire crews, including shift bosses, attended meetings; groups included from 12 to 20 individuals. The team-building sessions were task-centered; they focused on surfacing and solving problems that got in the way of working effectively, efficiently, and safely. The second author facilitated meetings as an outside consultant and helped the crews to bring problems related to their task to the surface. These groups identified task-related problems, generated ideas concerning the causes of the problems, solved problems and planned actions, and implemented and evaluated plans.

It should be noted that the team-building manipulation in this study was a global intervention. By definition, team building works with problems that teams identify. The specific problems addressed differed somewhat among the four crews. However, the basic process was the same; all team-building activities contained elements of the operational definition. They were (1) planned, (2) facilitated by an outside consultant, (3) engaged in by intact workgroups, (4) designed to develop the problem-solving capability of the crews, and (5) intended to solve major problems.

Description of specific team-building activities. Team building for each crew began with the consultant providing information on the background

---

[1] To "muck" means to remove the broken rock.

and purpose of the meetings. It was emphasized that the meetings' primary purpose was to improve the efficiency and safety of underground operations and to help make the mine a better place in which to work. Participants were told that all information gathered in the meetings would be used to generate constructive ideas for improvements and that their active participation was voluntary but welcome. The consultant then raised three specific questions: (1) How can we do the job better? (2) How can we make this a better place to work? (3) How can we make this a safer place to work? Focusing on these questions, crew members interacted with their shift bosses, with one another, and with the consultant. The consultant publicly recorded major points from the discussions on flip charts. He also condensed the problems, issues, and solutions from each meeting into major themes and fed them back to the crews via handouts at the beginning of each subsequent meeting. The mine's management also got a record of major themes for diagnosis of problems and planning of action.

Four major issues emerged from the team-building sessions: (1) poor communication between opposite shift crews and between miners and mine management; (2) poor working conditions underground in terms of both safety and performance; (3) poor repair and maintenance of equipment; and (4) pay and bid systems that were perceived as unfair. The crews' discussions generated a variety of suggestions for addressing these problems. The crews implemented some solutions directly. For example, in an effort to improve communications between opposite shifts, they scheduled daily meetings between opposite shift crews; none had existed before. The consultant facilitated several of these meetings between crews. Crews also acted on some suggestions involving underground working conditions, such as storing timber in safer, more accessible areas, cleaning up misplaced tools and materials, improving repair of tracks, and taking better care of drilling equipment.

During the fifth of the series of team-building meetings, the mine manager and superintendent interacted with the crews on issues that directly involved management of the mine and were beyond the decision making scope of the crews. Some of the problems were acted on immediately. For example, management freed some resources to create additional storage space underground. Some problems, primarily those dealing with maintenance and the pay system, were acknowledged as important and placed under study. Other issues, most notably the need for larger facilities and parking areas, were postponed for later resolution, since the mine had plans for future expansion. As management subsequently acted on problems, they made those actions known to the crews via the meetings and notices posted on a bulletin board.

Goal setting. This variable was operationally defined as the setting of specific, difficult, but attainable production goals and the receiving of weekly feedback on performance in relation to those goals. Performance feedback was an intergral part of the manipulation because previous research has found feedback to be a necessary condition if goals are to influence performance (Locke et al., 1981).

The goal-setting induction was implemented as follows. Two weeks prior to actual goal setting with stope miners, those shift bosses with miners assigned to the goal-setting condition received one hour of training, conducted by the first author, on setting goals and giving feedback on performance. Next, the goal-setting induction occurred in a one-hour meeting, also conducted by the first author, in which stope miners and shift bosses received graphs depicting the past 18 months of production in their stopes. They were asked to set specific, difficult, but attainable goals for tons per manshift and grade of ore[2] for the next three months. Goals were set participatively, with the miners and their bosses establishing mutually agreed upon goals that met the stated criteria. After goals were set the researchers recorded them on a feedback sheet and gave copies to each miner and shift boss. The feedback sheets, which were updated each week, included (1) graphs that displayed the recent output of that individual's stope in terms of tons per manshift and grade of ore, and (2) a statement that gave the goals that had been agreed on for the three-month trial period and described the particular stope's average output to date. Shift bosses were instructed to present feedback sheets to miners in the goal-setting condition weekly. Direct observations and post-experimental interviews confirmed that shift bosses did provide this regular feedback.

Dependent Variables

Two productivity measures were generated, one for quantity of production and one for quality of production. Tons per manshift is a standard measure of productivity in mines. It measures the amount of rock broken and removed from a stope per person per eight-hour shift. We used this variable, which was available from mine records every two weeks, as a measure of quantity of production. During the eight-month experimental period, 16 measures of tons per manshift were taken; the first two months of the period provided baseline data.

Grade of ore is also a standard measure of productivity in mines. It measures the numbers of ounces of silver and lead per ton of ore taken out of a stope. This variable, available from mine records monthly, was used as a measure of the quality of production. Eight measures of grade of ore were taken during the experimental period, including two measurements of baseline data.

A third dependent variable was strategy development, operationally defined as changes in miners' behaviors related to improving grade of ore. A document known as the grade strategy checklist provided a measure of grades. It listed seven specific observable behaviors related to grade improvement that stope miners could control. The mine geologist and the first author developed the checklist, which contained the following items: controlling the height and width of the stope, mining with the ore vein, putting in sufficient ground support, using proper slushing techniques, cleaning out rock carefully, and accurately locating the brow for a new cut. The mine geologist, who was blind to experimental condition, completed the checklist twice for each stope—once before the first goal-setting induction and once three months later. He rated each behavior according to the following scheme: good = 15; average = 10; below average = 5. A total of 105 points was a perfect score. The grade strategy checklist represented an unobtrusive measure of the extent to which miners were engaging in appropriate behaviors to improve grade of ore. It further served as a direct test of the hypothesis that goal setting and team building lead to the development of strategies.

---

[2] These two measures are defined in the section on dependent variables.

Measures of goal acceptance and goal difficulty were obtained after the goal-setting inductions to check for any differences among conditions on these variables. Following Latham and Steele (1983), we measured goal acceptance and goal difficulty with one face-valid question each; participants responded to these questions on a 5-point Likert-type format.

RESULTS

Productivity

Table 1 summarizes the means for all three productivity variables before and after the experimental manipulations. Figures 2, 3, and 4 display plots of the dependent measures over time.

TABLE 1
Summary of Means for Tons per Manshift, Grade of Silver, and Grade of Lead before and after the Interventions

Dependent Variables	Experimental Conditions	Before Interventions[a]		After Interventions[b]	
		Means	s.d.	Means	s.d.
Tons per manshift	Team building and goal setting	14.52	5.63	13.73	5.41
	Goal setting	19.32	8.28	23.16	12.51
	Team building	13.23	2.28	13.22	3.41
	Control group	8.38	3.43	8.91	3.91
Grade of silver (ounce of silver per ton of ore)	Team building and goal setting	11.08	1.43	15.24	4.01
	Goal setting	19.95	5.80	20.20	5.85
	Team building	10.75	1.06	14.30	0.45
	Control group	16.07	0.47	32.42	5.97
Grade of lead (percent of lead per ton of ore)	Team building and goal setting	7.36	1.01	7.58	1.05
	Goal setting	6.31	0.74	7.13	2.00
	Team building	6.98	0.64	9.33	0.79
	Control group	5.92	1.48	4.37	1.30

[a] Means shown are averages of the dependent measures taken during the three months preceding the interventions.

[b] Means shown are the averages of the dependent measures taken during the three months following the interventions.

FIGURE 2
Tons per Manshift Plotted over Time by Experimental Condition

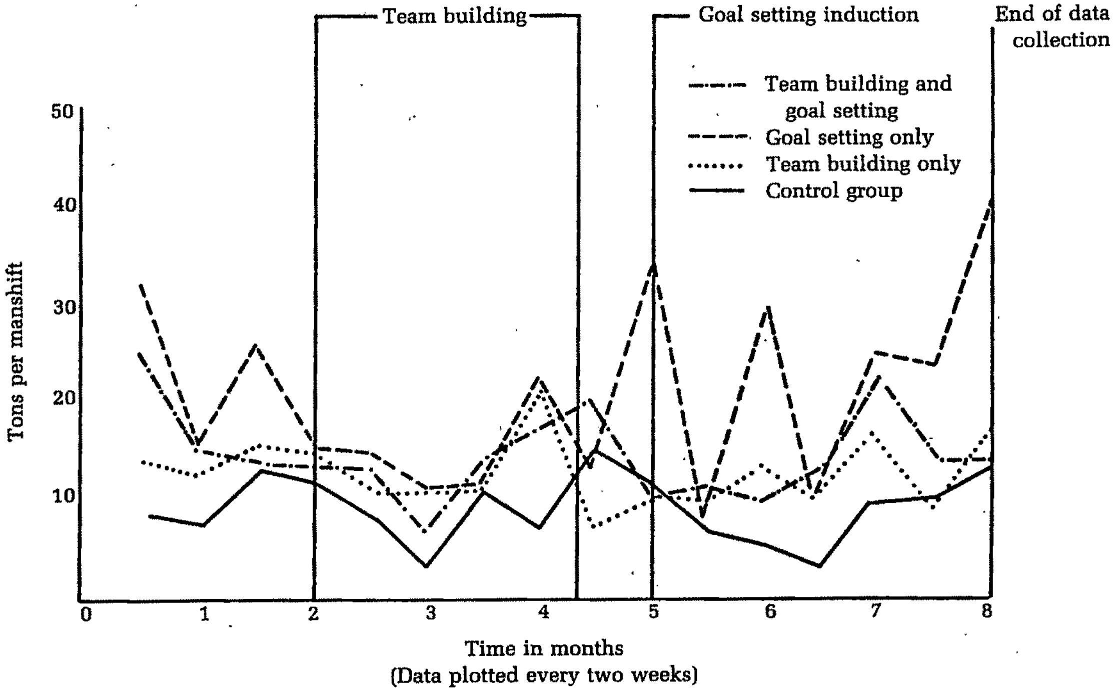


FIGURE 3
Grade of Silver Plotted over Time by Experimental Condition

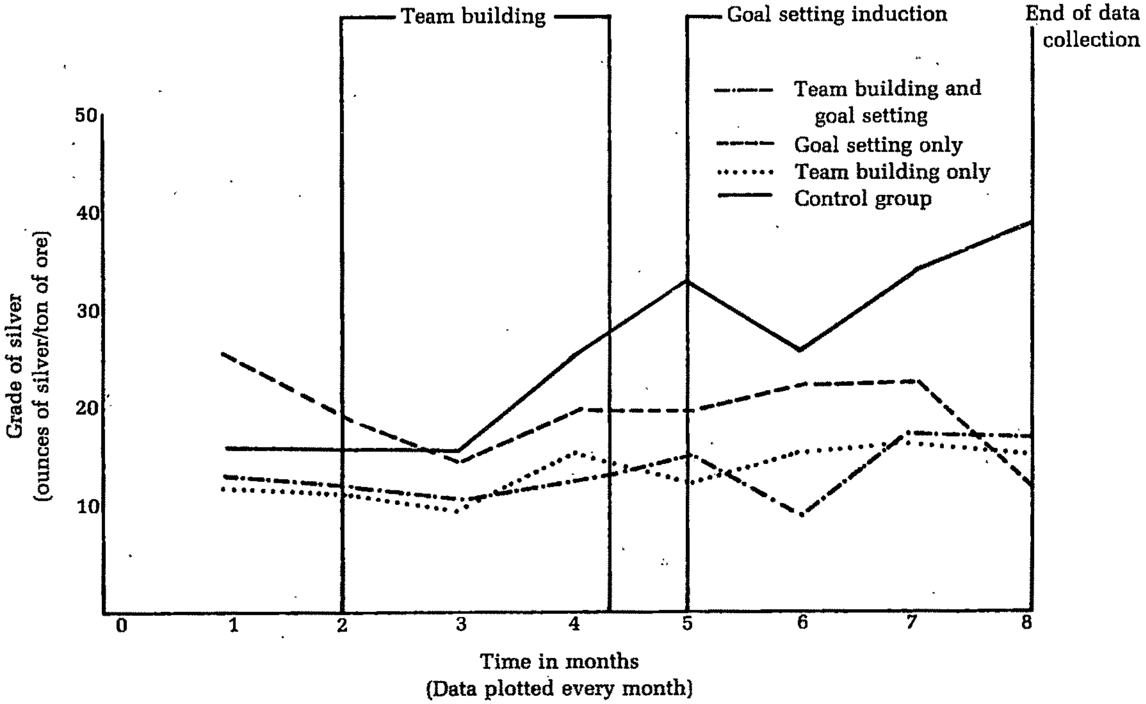
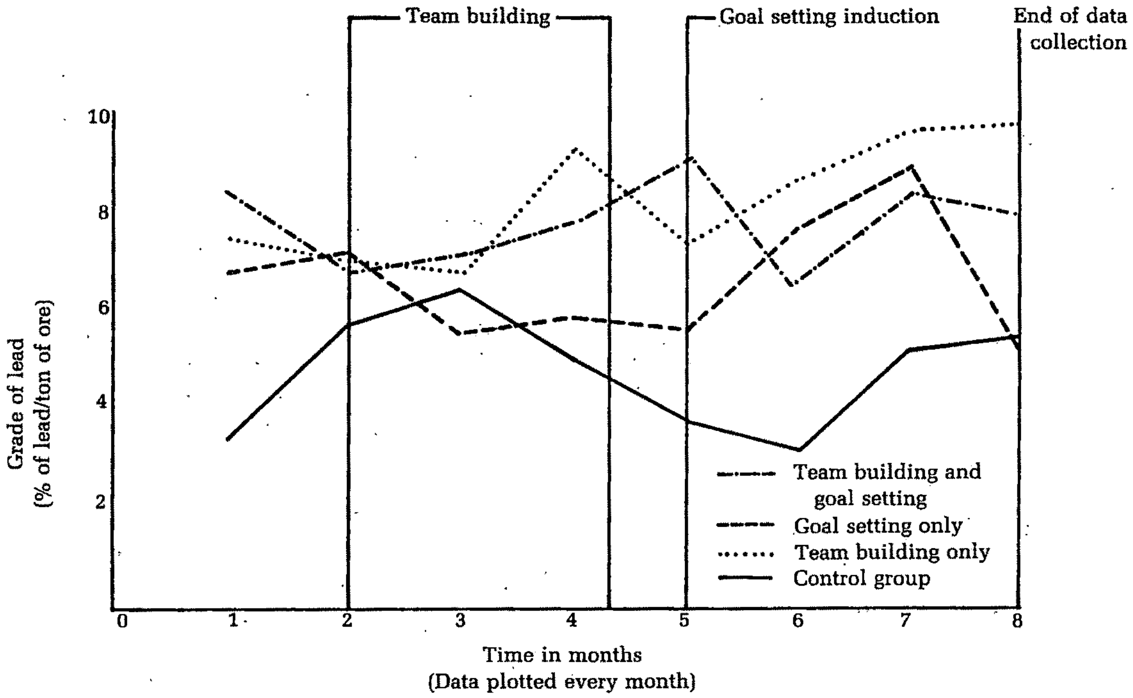


FIGURE 4
Grade of Lead Plotted over Time by Experimental Condition



Note that in Table 1 the largest increase in the mean grade of silver was in the control group. However, one stope accounted for 88 percent of this increase. In examining these results, the mine geologist stated that the large increase in grade in this stope resulted from a change in geological conditions rather than from improved performance of miners. We therefore conducted all of the following analyses with and without the data from this stope but found no appreciable differences between the results of the two analyses.

Since complete random assignment to experimental conditions was not possible, we followed procedures for analyzing quasi-experimental data suggested by Cook and Campbell (1979). First, we performed an analysis of variance on the productivity measures for the two-month baseline period preceding the experimental manipulations to determine if any significant differences existed among miners in the various conditions. There were no significant differences between conditions on pretest measures of tons per manshift and grade of lead, but there was a significant main effect for team building on grade of silver ($F_{1,16} = 4.28$, $p < .05$). In addition, we conducted an analysis of variance on measures of goal acceptance and goal difficulty to determine if the experimental conditions differed on these variables. No significant differences were found.

Next, an analysis of variance with repeated measures was performed for each of the three productivity variables using measures covering the eight-month experimental period. These results are given in Table 2. They show that there was a significant difference between stopes in the goal-setting and no-goal-setting conditions on tons per manshift but not on grades of silver or lead. There was a marginally significant difference between the team-building and no-team-building conditions on grade of silver but not on tons per manshift or grade of lead. There was a marginally significant interaction between team building and goal setting on tons per manshift. Additionally, there was a significant time effect of the two interventions on tons per manshift and a marginally significant time effect on grade of silver. Finally, we found a marginally significant team × goal × time interaction on grade of silver. We found no significant differences among any of the conditions on grade of lead.

The results of the analysis of variance with repeated measures are, unfortunately, somewhat equivocal. We found no significant difference among conditions on preexperimental measures of tons per manshift, but found a significant difference between the goal-setting and no-goal-setting conditions over the course of the experiment. This finding, in conjunction with the results summarized in Figure 2, provides some support for a positive effect for goal setting. Yet, we cannot draw an unambiguous conclusion for two reasons. First, even though the difference before the experiment between the goal-setting and no-goal setting conditions on tons per manshift was not significant, it was substantial and could by itself conceivably account for the results. Second, the small sizes of cells along with considerable variability in dependent measures could have produced spurious results.

TABLE 2
Analyses of Variance with Repeated Measures for Evaluating the Effects of Team Building and Goal Setting on Tons Per Manshift[a] and Grade of Silver[b]

	Tons per Manshift					Grade of Silver				
Sources	*df*	*MS*	*F*	% Variance	*p*	*df*	*MS*	*F*	% Variance	*p*
Goals	1	3684.10	7.1	44.5	.02	1	323.16	.4	2.3	n.s.
Team	1	86.96	.2	1.0	n.s.	1	3775.20	4.2	26.4	.06
Team × goals	1	1729.76	3.3	20.9	.09	1	320.15	.4	2.2	n.s.
Residual	16	517.41				16	892.89			
Time	15	300.36	1.8	11.5	.03	7	161.13	1.8	11.4	.09
Goals × time	15	94.25	.6	3.6	n.s.	7	130.91	1.5	10.0	n.s.
Team × time	15	182.91	1.1	6.9	n.s.	7	45.65	.5	3.2	n.s.
Team × goals × time	15	205.65	1.3	7.9	n.s.	7	172.13	1.9	12.1	.07
Residual	240	163.37				112	88.54			

[a] This analysis includes tons per manshift measures for 16 time periods covering the eight-month experimental period.

[b] This analysis includes grade of silver measures for 8 time periods covering the eight-month experimental period.

The meaning of the marginally significant difference between the team-building and no-team-building conditions on grade of silver is also unclear. There was a significant difference between these conditions on grade of silver before the intervention, with miners assigned to the no-team-building condition apparently outperforming those assigned to the team-building condition. This initial difference between the two conditions did not change during the course of the experiment. However, we cannot conclude that the absence of team building led to greater improvement of performance than the presence of team building. In addition to small cell sizes and high variability in the data, the lack of random assignment to team-building conditions prohibits firm conclusions. It is possible that other factors, such as history, selection-maturation interactions, or other threats to internal validity caused the difference found between conditions (Campbell & Stanley, 1963).

Next, an analysis of variance with gain scores was conducted. This procedure, suggested by Reichardt (1979), examines mean changes in performance between treatment groups from pretest to posttest. It is a useful technique for analyzing data from nonequivalent group designs because, like analysis of covariance, it helps to control for any pretest differences between groups. Analysis of gain scores was more appropriate than ANCOVA with our data since the pretest differences between treatment conditions were substantial. Unlike the ANCOVA method, the precision of analysis of gain scores is not related to the size of pretest differences (Reichardt, 1979). The procedure employed a 2 × 2 analysis of variance with goal setting and team building as independent variables and scores for changes in productivity as the dependent variables. Changes were measured both as absolutes and as percentages. We conducted the analysis both with and without data from the unusual stope in the control group. No significant main or interaction effects were found, indicating that the treatments had no effect. But again we cannot draw firm conclusions because, as Reichardt (1979) pointed out, analysis of gain scores is subject to other problems—bias from selection-maturation interactions, floor and ceiling effects, and regression effects—threats to validity that simply cannot be ruled out.

To analyze the effects of treatments further, we assessed changes within each experimental condition before and after the inductions. Using time as an independent variable with two levels—average performance measures both before and after interventions—we conducted a one-way analysis of variance on tons per manshift, grade of silver, and grade of lead for each of the four experimental conditions. The results showed that the goal-setting-only condition had a marginally significant effect on tons per manshift ($F_{1,3} = 7.1$, $p<.08$); results for no other conditions approached significance. The combined team building–goal setting condition had a marginally significant effect on grade of silver ($F_{1,5} = 4.5$, $p<.09$) and so did the team-building-only condition ($F_{1,5} = 4.9$, $p<.08$). The team-building-only condition had a marginally significant effect on grade of lead ($F_{1,5} = 4.7$, $p<.08$), with no other conditions approaching significance. Although the results of these within-group analyses are only marginally significant, they provide some evidence for some treatment effects. Specifically, team building showed a positive effect on grades of silver and lead, and goal setting appeared to improve tons per manshift. Again, since the treatment groups were nonequivalent and the cell sizes small, we cannot entirely rule out factors other than the treatments as explanations of the results obtained.

In summary, results of analyses of the data on productivity were not conclusive. The weight of evidence, however, suggests that the treatments produced little or no effect on performance. Although there is some evidence consistent with a positive but small effect for goal setting on tons per manshift, rival explanations of this result could not be ruled out entirely. There was also some evidence for a positive effect for team building on grade of ore, but this evidence is open to even more questions.

Strategy Development

An analysis of variance on pretest scores on the grade strategy checklists showed no significant differences among experimental conditions. An analysis of variance with repeated measures was then performed on pretest and posttest scores on grade strategy checklists. We found no significant main effects for either team building or goal setting. There was, however, a main effect for time ($F_{1,16} = 19.7$, $p<.001$), indicating that miners improved grade strategy across conditions over time. When we submitted scores from grade strategy checklists to an ANOVA with the four experimental conditions as separate treatments and time as the independent variable, there was a significant effect over time for the team-building-only condition ($F_{1,5} = 14.3$, $p<.01$), and marginally significant effects for the goal-setting-only ($F_{1,3} = 7.7$, $p<.07$), and the combined team building-goal setting condition ($F_{1,5} = 4.8$, $p<.08$). The control group showed no significant improvement on scores for grade strategy. This within-group analysis suggests that the major gains in strategy development over time were achieved in the experimental treatment conditions. However, when an analysis of variance using gain scores was conducted, results showed no significant differences among conditions. Table 3 summarizes the mean scores on the grade strategy checklist for experimental conditions before and after the experimental manipulations.

Qualitative Assessment of the Interventions

Although it was not the primary emphasis of this study, we also assessed participants' reactions to the interventions. After the interventions, the first author conducted structured interviews with three randomly selected shift bosses and ten randomly selected hourly employees; these interviews provided data for this evaluation. In addition, these individuals completed a brief postexperimental questionnaire pertaining to the specific problems identified by hourly employees in the team-building sessions.

In general, the structured interviews revealed that miners liked both interventions. With respect to team building, most miners thought that the intervention allowed them to express their ideas to management and get

TABLE 3
Summary of Mean Scores on the Grade Strategy Checklist[a] before and after the Interventions

Experimental Conditions	Before Interventions	After Interventions
Team building and goal setting	55.00	70.00
Goal setting	70.00	85.00
Team building	65.83	80.00
Control group	62.50	66.25

[a] Maximum possible score on the grade strategy checklist was 105.

things out in the open. Several employees observed that there seemed to be better relations between management and the union as a result of the intervention. Specific comments included: "The meetings brought a lot of stuff out, got problems out into the open," "The guys felt better identifying problems," and "The meetings helped us to talk better with the opposite shift." But several hourly employees mentioned a shortcoming that team building had: nothing seemed to change much as a result of the meetings. Others noted a lack of interest on the part of the hourly employees and that it was a nuisance to come to the meetings. Some representative comments were: "Nothing really happened as a result of the meetings," and "I'd like to see the people take a bigger part; if more would have talked up, things would have been better, more things would have changed."

Those miners involved in the goal-setting intervention most often mentioned regular feedback on their productivity and the grades they were obtaining as a valuable part of the program. Representative statements were: "Goal setting made us more conscious of what we are doing," and "It showed me how I can make a little better money—set standards for myself." However, two miners expressed suspicion regarding goal setting: "I thought that the goals would be used to set day's pay and incentive rates."

As noted in the Methods section, the team-building sessions identified several specific problems: communication between opposite shift crews; communication among underground management personnel; politics and favoritism in pay and bid systems; equipment and maintenance; and working and mine conditions. We administered a brief postexperimental questionnaire to determine these shift bosses' and employees' perceptions of the effects of team building on these problems. The questionnaire simply asked: "For each of the following problems listed, please indicate if the problem is ________ as a result of the crew meetings"; the response format ranged from 1 = much worse to 5 = much better. Results indicated that workers perceived improvement in performance in all of the problem areas. The largest perceived improvement was in communication between opposite shift crews ($\bar{x}$ = 4.07), followed by politics and favoritism ($\bar{x}$ = 3.92), communications among underground management personnel ($\bar{x}$ = 3.85), equipment and maintenance ($\bar{x}$ = 3.46), and working and mine conditions ($\bar{x}$ = 3.23).

To summarize, the data obtained through structured interviews indicated that shift bosses and hourly employees generally liked and accepted the interventions, although we noted some negative reactions. Further, results from the questionnaire showed that employees perceived team building as having improved the specific problems it had identified. Obviously, there are problems with the validity and reliability of these qualitative findings. However, they provide some additional evidence consistent with positive outcomes of the interventions.

DISCUSSION

Although our design was an improvement over past research in many respects, the results of this field experiment were not conclusive. Even when results are in the predicted direction, we cannot say with certainty that goal setting or team building produced these effects in performance or strategy development. The shortcomings of this experiment illustrate the many difficulties encountered in doing field research.

The main problem in interpreting our results is the fact that we could not ensure that experimental and control groups were equivalent. Differences between groups owing to selection, maturation, or instrumentation could therefore have influenced the findings. We tried several, but cannot be certain that *any* statistical procedure we employed accounted for all potential biases in our results. As Cochran and Rubin (1973) pointed out, "If randomization is absent, it is virtually impossible in many practical circumstances to be convinced that the estimates of the treatments are in fact unbiased" (1973: 417). Cook and Campbell (1979) suggested that, in interpreting the results of such quasi experiments, it is important to consider their consistency with previous research. The findings of previous research on goal setting are quite clear: hard, specific goals, if accepted, lead to improved task performance. The results of this experiment, although in the positive direction, are not as clear or dramatic as those of previous studies.

Several factors relating to the research design and setting may help to explain our relatively weak results. First, the small number of stopes in the analysis was a limitation, as was the high degree of variability in the performance data. Both limitations make it difficult to detect systematic variance. In addition, the experiment was conducted with unionized hard-rock miners—subjects who were skeptical of modern management techniques. As two employees remarked in the postexperimental interviews, there was some concern that the goal-setting intervention was an attempt by management to boost productivity and then adjust the incentive pay scales downward. If this sentiment was widespread, it may have caused some miners to restrict their output. Although miners in the goal-setting condition indicated a high level of goal acceptance, we cannot be certain that this acceptance was real because of problems with the measure of acceptance used in the study. Specifically, our measure of goal acceptance consisted of a single, face-valid question whose validity has been questioned because goal-setting participants tend to respond positively to it with little variance in their responses (Locke et al., 1981). And, even though employees expressed generally positive reactions to the goal-setting intervention in the structured interviews, reactive effects of the interview itself may have colored their responses. Thus, the weak results may simply reflect miners' relatively weak acceptance of or commitment to the goals set. Finally, the nature of the miners' task and working environment may have contributed to the minimal results. Hard-rock mining is a difficult task in an uncertain environment. Although the performance of miners is a function of their ability and motivation just as it is in other jobs, events beyond their control can influence their output, perhaps more so than in many other blue-collar jobs. In such an environment, it is possible that increases in miners' efforts and strategy development did not translate into improved performance.

Yet, these limitations of the study are also its strengths. Previous research on goal setting has been conducted primarily in laboratories under controlled conditions. Those field experiments that have been conducted have generally had nonunionized participants performing relatively simple, routine tasks. As others have pointed out, problems may arise in applying goal setting in situations in which tasks are complex, unstructured, or interdependent (Baumler, 1971; Latham & Yukl, 1975; Locke & Latham, 1984). The present study suggests that the effects of goal setting on performance may be muted under such conditions. Further research in similar settings is needed to continue to assess the degree to which goal-setting effects can be generalized.

Also, previous research examining the effects of team building on performance has been inconclusive, largely because of poor research designs and reliance on subjective measures. The present study, employing a more rigorous design and objective measures, also produced ambiguous results. Although a modest improvement in grade of ore occurred in the team-building conditions, we cannot be certain why it occurred. Several factors, in addition to those cited above, may have contributed to the results obtained. First, the team-building intervention was more diffuse than the goal-setting intervention. Although increased effectiveness in production was one of the objectives of team building, the sessions did not focus specifically on increasing tons per manshift and quality of ore. The broad nature of this intervention could have reduced its impact on these output variables, even though miners perceived improvements on the specific problems identified by team building. Related to this issue is the fact that we evaluated performance for only three-and-a-half months following the intervention. Because team building involves rather complex and relatively slow processes of developmental change (French & Bell, 1984; Porras & Berg, 1978), our analysis may not have captured the full effects of the intervention. A longitudinal assessment is now being conducted to examine long-term effects.

Two final problems in interpreting the results on productivity concern possible selectivity bias and reactive effects of the experiment. The mine superintendent made the nonrandom assignment of stopes to the team-building condition primarily because those stopes were on the two major production levels of the mine. And, although nothing suggests that miners in the team-building condition differed from those in the non-team-building condition, the geological conditions of the stopes differed between these treatment groups. Stopes in the non-team-building group were primarily at high elevations in the mine, where the rock is harder than at low elevations. Generally, the harder the rock, the faster miners can drill, blast, and muck. The non-team-building stopes also apparently had richer veins. These different geological conditions may explain the preexperimental differences between treatment groups. The geological differences between stopes were impossible to control for and make it difficult to compare effects on performance between treatment groups. Under these conditions, the results of the analyses within treatment groups, which examined changes in performance over time, are the most informative.

The reactive effects of the experiment may also have caused problems. As mentioned, employees in this experiment were not used to being studied by people other than "efficiency experts." Although there is no strong evidence, some miners may have suppressed their output out of suspicion of the researchers' motives. On the other hand, the modest performance effects found may have been due to miners' increased motivation resulting from the so-called Hawthorne effect. The structured interviews and informal observations revealed that many miners felt positive about, and even enjoyed, all the attention from the researchers and mine management. These possible Hawthorne effects do not overly concern us because one of the causal elements of most behavioral science approaches to improving productivity, particularly an approach like team building, is increased attention and focus on employees and their work. In fact, we argue that the objective of many such interventions is a *sustained* Hawthorne effect.

Finally, we cannot entirely rule out the possibility that competition between groups or diffusion of the treatments between groups affected the results. The interventions were public in the sense that employees knew who received team building or goal setting. Although there is no evidence to suggest this, it is possible that rivalry could have emerged between the haves and the have-nots. Diffusion of the interventions also could have occurred, but this was not very likely, particularly in the team-building condition. Miners in the control group could have set personal goals as a result of the intervention, but there is no evidence that they did. Only those miners in the goal-setting condition received weekly performance-feedback graphs on tons per manshift and grade of ore to minimize possible contamination of the no-goal-setting groups. However, all miners received bimonthly information on tons per manshift on their paychecks; they did not, however, get feedback on grade of ore. Thus, miners in the goal-setting condition received more frequent and specific feedback than other miners.

The present research produced some information, consistent with previous research, regarding the cognitive effects of goal setting. Miners who set goals and received periodic feedback appeared to develop and apply better strategies for improving the grade of ore they produced than miners in the control group. Again, the limitations cited for the productivity variables make firm conclusions unwarranted. It is also important to note that any improvements in grade strategy owing to the effects of goal setting were not translated into increased grade quality. The weak effects of strategy development or conditions beyond the miners' control affecting grade of ore may have been responsible for this lack. The study suggests, however, that cognitive mechanisms as well as motivational mechanisms can affect the relationship between goal setting and task performance. Employees who set specific hard goals and receive feedback on them may not only work harder than employees who do not set goals and receive feedback—they may also develop more appropriate ways to accomplish task goals. Team building also appeared to have a positive effect on strategy development, consistent with Hackman's (1976) model. And, even though strong conclusions cannot be drawn, it is interesting to note that team building may have influenced the development of strategies which in turn may have improved quality of grade.

In summary, although the findings of this field experiment are unclear, the results are important for several reasons. First, the study raises a question about the extent to which goal-setting effects are generalizable to union employees working on complex and uncertain tasks. Further research under similar conditions is warranted. Second, the study suggests that goal setting and team building may affect cognitive mechanisms such as strategy development. Future research should continue to examine this mechanism using more refined measures. Third, the study points to our continuing lack of understanding of how team building influences performance. We attempted to operationally define this intervention more clearly and to examine its effects more systematically than previous research has done, and yet its effects remain elusive. Further research should attempt to document team-building activities and possible intervening mechanisms more precisely and examine their relation to outcome variables. Perhaps path analysis or other causal analysis techniques (James, Muliak, & Brett, 1982) might provide useful means for assessing the multiple relationships generated by the team-building intervention. Finally, declining productivity continues to be a major problem in underground mines and other organizations, and managers continue to be beseiged with a large number of quick-fix solutions from behavioral scientists and others. Unfortunately, many approaches to improving employees' productivity remain untested. The present research, by providing a systematic evaluation of two specific techniques for improving the performance of employees, provides managers with a more objective basis on which to judge the usefulness of team building and goal setting.

REFERENCES

Bandura, A., & Simon, K.M. 1977. The role of proximal intentions in self-regulation of refractory behavior. ***Cognitive Therapy and Research,*** 1: 177–193.

Baumler, J.V. 1971. Defined criteria of performance in organizational control. ***Administrative Science Quarterly,*** 16: 340–350.

Bavelas, J.B., & Lee, E.S. 1978.Effects of goal level on performance: A trade-off of quantity and quality. ***Canadian Journal of Psychology,*** 32: 219–240.

Beer, M. 1976. The technology of organization development. In M.D. Dunnette (Ed.), ***Handbook of industrial and organizational psychology:*** 937–993. Chicago: Rand McNally & Co.

Buller, P.F. 1982. ***The effects of team building–problem solving and goal setting–performance feedback on the productivity of underground metal miners: A field experiment.*** Unpublished doctoral dissertation, University of Washington, Seattle.

Buller, P.F. 1983. Team building-problem solving and performance: Conceptual issues and a theoretical model. In D. F. Ray (Ed.), ***Applying systematic knowledge to problems of organization and management:*** 288–290. Mississippi State University: Southern Management Association.

Campbell, D.T., & Stanley, J.C. 1963. ***Experimental and quasi-experimental designs for research.*** Chicago: Rand McNally & Co.

Christensen-Szalanski, J.J.J. 1980. A further examination of the selection of problem-solving strategies: The effects of deadlines and analytic aptitudes. ***Organizational Behavior and Human Performance,*** 25: 107–122.

Cochran, W.G., & Rubin, D.B. 1973. Controlling bias in observational studies: A review. ***Sankhya,*** 35: 417–446.

Cook, T.D., & Campbell, D.T. 1979. ***Quasi-experimentation: Design and analysis issues for field settings.*** Boston: Houghton-Mifflin.

Cummings, T.G., Molloy, E.S., & Glen, R.H. 1974. Intervention strategies for improving productivity and the quality of work life. ***Organizational Dynamics,*** 4: 52–68.

DeMeuse, K.P., & Liebowitz, S.J. 1981. An empirical analysis of team building research. ***Group and Organizational Studies,*** 6: 357–378.

Dyer, W.G. 1977. ***Team building: Issues and alternatives.*** Reading, Mass.: Addison-Wesley.

French, W.D., & Bell, C.H. 1984. ***Organization development: Behavioral science interventions for organization improvement*** (3rd ed.). Englewood Cliffs, N.J.: Prentice-Hall.

Hackman, J.R. 1976. Group influences on individuals in organizations. In M.D. Dunnette (Ed.), ***Handbook of industrial and organizational psychology:*** 1455–1525. Chicago: Rand McNally & Co.

Huse, E.F. 1980. ***Organization development and change*** (2nd ed.). St. Paul, Minn.: West Publishing Co.

James, L.R., Muliak, S.A., & Brett, J.M. 1982. ***Causal analysis: Assumptions, models, and data.*** Beverly Hills, Calif.: Sage Publications.

Latham, G.P., & Baldes, J.J. 1975. The "practical significance" of Locke's theory of goal setting. ***Journal of Applied Psychology,*** 60: 122–124.

Latham, G.P., & Steele, T.P. 1983. The motivational effects of participation versus goal setting on performance. ***Academy of Management Journal,*** 26: 406–417.

Latham, G.P., & Yukl, G.A. 1975. A review of research on the application of goal setting in organizations. ***Academy of Management Journal,*** 18: 824–845.

Locke, E.A. 1968. Toward a theory of task motivation and incentives. ***Organizational Behavior and Human Performance,*** 3: 157–189.

Locke, E.A., Feren, D.B., McCaleb, V.M., Shaw, K.N., & Denny, A.I. 1980. The relative effectiveness of four methods of motivating employee performance. In K.D. Duncan, M.M. Gruneberg, & D. Wallis (Eds.), ***Changes in working life***: 363–388. New York: John Wiley & Sons.

Locke, E.A., & Latham, G.A. 1984. ***Goal setting: A motivational technique that works.*** Englewood Cliffs, N.J.: Prentice-Hall.

Locke, E.A., Shaw, K.N., Saari, L.M., & Latham, G.P. 1981. Goal setting and task performance: 1969-1980. ***Psychological Bulletin,*** 90: 125–152.

Margulies, N., & Raia, A.P. 1968. People in organizations: A case for team training. ***Training and Development Journal,*** 22(8): 2–11.

Naylor, J.C., Pritchard, R.D., & Ilgen, D.R. 1980. ***A theory of behavior in organizations.*** New York: Academic Press.

Nicholas, J.M. 1982. The comparative impact of organization development interventions on hard criteria measures. ***Academy of Management Review,*** 7: 531–542.

Porras, J.I., & Berg, P.O. 1978. Impact of organization development. ***Academy of Management Review,*** 3: 249–266.

Reichardt, C.S. 1979. The statistical analysis of data from nonequivalent group designs. In T.D. Cook & D.T. Campbell (Eds.), ***Quasi-experimentation: Design and analysis issues for field settings:*** 147–205. Boston: Houghton-Mifflin.

Shaw, K.N. 1983. ***The relationships among goals, strategies, and task performance.*** Working paper, University of Kentucky, Lexington.

Taylor, F.W. [1911] 1967. ***The principles of scientific management.*** (Reprint) New York: Norton.

Terborg, J.R. 1976. The motivational components of goal setting. ***Journal of Applied Psychology,*** 61: 613–621.

Terborg, J.R., & Miller, H.E. 1978. Motivation, behavior, and performance: A closer examination of goal setting and monetary incentives. ***Journal of Applied Psychology,*** 63: 29–39.

Woodman, R.W., & Sherwood, J.J. 1980. The role of team development in organizational effectiveness: A critical review. ***Psychological Bulletin,*** 88: 166–186.

Paul Buller earned his Ph.D. degree at the University of Washington; he is an assistant professor of management in the Department of Business Administration at Utah State University. His current research interests include productivity improvement, organizational change, and strategic management.

Cecil Bell earned his Ph.D. degree at Boston University; he is an associate professor of management and organization in the Graduate School of Business Administration at the University of Washington. His current research interest is organization development.

© *Academy of Management Journal*
1986, Vol. 29, No. 2, 329–346.

ENVIRONMENTS AS MODERATORS OF THE RELATIONSHIP BETWEEN STRATEGY AND PERFORMANCE

JOHN E. PRESCOTT
University of Pittsburgh

Environments have been viewed as key contingency variables for relationships between strategy and performance. Yet, the nature of this contingency relationship has not been resolved; do environments moderate its form or its strength? Using moderated regression analysis and subgroup analysis, this study demonstrated that environments, as measured by characteristics of market structures, moderate the strength but not the form of relationships between strategy variables and performance.

Both a business' environment and its strategy have been hypothesized and empirically demonstrated to have significant effects on performance (Porter, 1980; Scherer, 1980). Previous research has considered strategy to be basically under the control of managers, but has viewed environments as constraints that in certain situations managers can proactively change (Hofer & Schendel, 1978; Pfeffer & Salancik, 1978). Thus, much of the strategic management literature has focused on the relationship between strategy and performance and considered environments as moderators of that relationship. The field of industrial organizational economics has emphasized the linkage between environment and performance and thus has viewed environments as primary determinants of performance (Porter, 1981). Recently, there have been attempts to examine the relationships among environment, strategy, and performance variables (Hambrick, 1986; Hitt, Ireland, & Stadter, 1982; Jauch, Osborn, & Gluck, 1980). However, research examining that threefold relationship has not adequately addressed the issue of whether environments are (1) independently related to performance, (2) moderators of the relationship between strategy and performance, or (3) some combination of the two. The purpose of the present study is to examine this issue.

The research question was this: do environments independently influence performance, or do environments modify the strength or the form of the relationship between strategy and performance? This question addresses the heart of contingency theory (Schoonhoven, 1981). If environments independently influence performance, then contingency relationships do not exist. If

---

The author gratefully acknowledges the generous support from the Strategic Planning Institute, Cambridge, Massachusetts. The author benefited greatly from the comments and assistance of Ari Ginsberg during the project.

environments modify the strength of the relationship between strategy and performance, the role of contingency theory would be to identify meaningful subenvironments and to determine the relative strength of the effects of these subenvironments on relationships between strategy and performance. If, on the other hand, environments modify the form of the relationship between strategy and performance, the role of contingency theory would be to identify key interactions and establish their links to performance.

THEORETICAL BACKGROUND AND PAST RESEARCH

Scholars in the field of strategic management have conceptualized environment as one of the key constructs for understanding organizational behavior and performance (Hofer & Schendel, 1978). Basing their research on an extensive review of over 80 articles, Lenz and Engledow (1986) identified five approaches to modeling environments: the industry structure model (Porter, 1980), the cognitive model (Weick, 1979), the organizational field model (Dill, 1958), the ecological and resource dependency model (Aldrich, 1979; Pfeffer & Salancik, 1978), and the era model (Naisbitt, 1982). All these approaches to modeling environments vary in terms of assumptions about environmental structures, assumptions about the process and causes of environmental change, and assumptions about how managers or researchers know and understand environments (Lenz & Engledow, 1986).

Regardless of how environments are modeled, research findings suggest that their characteristics influence decision making through managerial perceptions and objective dimensions of industries' structures. For example, Miles and Snow's (1978) research illustrated that managers enact their domains by focusing on certain conditions, trends, and events in their environments. Snow and Miles (1983) further suggested that "those factors that go unnoticed or are deliberately ignored are not part of the organization's enacted environment and thus do not affect managerial decision making and action" (1983: 239).

Lawrence and Lorsch (1967) found that managers in the standardized container industry perceived their environment to be more certain than managers in the plastics and consumer foods industries. Further, they found that the dominant environmental requirements varied among industries; product and technological development dominated in plastics, product development and market research in consumer foods, and customer service and product quality in standardized containers.

The basic model for industrial organizational economics is the structure-conduct-performance paradigm. According to the paradigm, the performance of a firm depends on its conduct in such matters as pricing policies, R & D emphasis, and investment policies. Conduct, in turn, depends on the structure of the industry, which includes the barriers to entry, level of concentration, and degree of product differentiation (Scherer, 1980).

Economists employing the market structure-conduct-performance paradigm have suggested for years that objective characteristics of industries such as rates of growth and concentration levels affect both the conduct or strategy of firms and their performance (Bain, 1959; Scherer, 1980). The Profit Impact of Market Strategy[1] (PIMS) studies found that industry characteristics have an effect on the performance of business units, and Rockart (1979) suggested that industry characteristics determine certain factors critical for the success of business strategies. Porter (1980) developed probably the most comprehensive treatment of industry influences on firms' strategies and performance levels. Overall, these findings have suggested that both the perceived and objective environments of an industry moderate the relationship between a business' strategy and its performance. That is, different environments will have different sets of critical success factors (Rockart, 1979). However, there have been very few attempts to examine the moderating effects of environments on the relationship between strategy and performance.

Moderating Effects of Environments

In general, studies can examine moderating effects in two ways: either (1) by including interaction variables in an additive model, or (2) by estimating parameters of an additive model for subgroups of a total sample (Arnold, 1982; Sharma, Durand, & Gur-Arie, 1981). Analysis using interaction terms examines the form of a relationship and subgroup analysis examines the strength of a relationship. Jauch, Osborn, and Glueck (1980), employing moderated regression analysis, examined the interacting effects of nine environmental changes on eight strategic decisions and two performance measures. Their results indicated that environmental change has little relationship to performance, although strategic decisions concerning financial changes and production efficiency were statistically significantly related to performance. The interaction of environmental changes and strategic decisions was not significant. The only other study found that used moderated regression analysis to examine relationships between environment, strategy, and performance was by Hitt, Ireland, and Stadter (1982). Their study examined separately the moderating effects of two variables, grand or corporate strategy and type of industry, on the strategic importance of functional areas and company performance. Their results indicated that both grand strategy and type of industry had separate, significant moderating effects on the relationship between functional areas and performance. An attempt to examine the three-way interaction of industry type and grand strategy simultaneously on the relationship between functional importance and performance was not possible because their samples were small. Thus, the question of whether grand strategy and industry type simultaneously have an interacting effect on the relationship between functional importance and performance was left unanswered.

---

[1] For a comprehensive review of research on strategy that has used the PIMS data base, see Ramanujam and Venkatraman (1984).

Prescott (1983) used analysis of variance to examine the effects of the interaction of Porter's (1980) sets of generic strategies and generic industries on the performance of business units. Porter's generic strategies include overall cost leadership, focus, and differentiation (1980: 34–46); his generic industries include fragmented, mature, global, declining, and emerging (1980: chaps. 9–14). Generic strategies explained 12 percent of variance in performance, and generic industries and the interaction of the two sets explained 1 and 2 percent respectively. A potential problem with using such generic categories is a loss of information through aggregation.

The second approach to examining the moderating effects of environments on the relationship between strategy and performance is subgroup analysis. The assumptions underlying subgroup analysis are that an interaction term of the moderated regression type is not significant and that the error term is posited to be a function of the moderator variable (Sharma et al., 1981). Therefore, if subgroups that are homogeneous in terms of error variance are formed from the total sample, the predictive efficacy of the subgroups should increase. The typical tests for the existence of subgroups are the chi-square test, the Chow test (Chow, 1960), and their extensions (Neter & Wasserman, 1974). Many studies have employed subgroup analysis in the context of the relationship between strategy and performance (Bass, 1974; Bass, Cattin, & Wittink, 1977; Gale, 1972; Hambrick, 1986; Hatten & Schendel, 1977; Hatten, Schendel, & Cooper, 1978; Newman, 1978; Oster, 1982; Phillips, Chang, & Buzzell, 1983; Porter, 1979; Prescott, 1983). Two major issues with which these studies have been concerned are the formation of subgroups and the testing of subgroups for homogeneity. The studies have generally concluded that use of subgroups increases the predictive efficacy of regression equations and that the strength of strategy variables varies across the subgroups.

In sum, although studies employing moderated regression analysis have examined the ways in which environments affect the form of the relationship between strategy and performance, studies employing subgroup analysis have typically examined only the strength of this relationship across various subgroups and assumed that there was no statistically significant interaction. No study has had the explicit purpose of examining both the form and the strength of the moderating effect of environments on the relationship between strategy and performance.

TYPOLOGY OF SPECIFICATION VARIABLES

Building on the work of two previous researchers (Lazarsfeld, 1955; Rosenberg, 1968), Sharma and his colleagues (1981) developed a typology of specification variables. A specification variable is one that specifies the form or the strength, or both, of the relationship between a predictor and a criterion. Figure 1 depicts this typology, which has two dimensions. Shown on the horizontal axis, the first dimension addresses whether a specification variable is related to a criterion variable. On the vertical axis, the second dimen-

FIGURE 1
Typology of Specification Variables[a,b]

	Environment related to criterion, predictor, or both	Environment not related to criterion or predictor
No significant interaction of environment with predictor	Intervening, Exogenous, Antecedent, Suppressor, Predictor Cell 1	Moderator (Homologizer) Cell 2
Significant interaction of environment with predictor	Moderator (Quasi moderator) Cell 3	Moderator (Pure moderator) Cell 4

[a] Note that in this study the predictor variables were strategy variables, the criterion variable was ROI, and the specification variable was environment.

[b] Adapted from Sharma, Durand, and Gur-Arie (1981).

sion addresses whether the specification variable interacts with the predictor variable. For the purposes of the present study, the predictor variables were strategy variables, the criterion variable was return on investment (ROI), and the specification variable was environment.

In Figure 1, the specification variable could be related to the criterion variable or to the predictor variable or to both without interacting with the predictor. This specification variable might then be called an intervening, exogeneous, antecedent, suppressor, or predictor variable, on the basis of other considerations. Cell 1 represents these possibilities. The remaining three cells represent different types of moderator variables. Cell 2 represents a moderator variable that is not significantly related to the criterion and predictor variables and does not interact with the predictor. This type of moderator variable, referred to as a homologizer, influences the strength of the relationship between the predictor and criterion variables across homogeneous subgroups (Arnold, 1982; Zedeck, 1971).

Moderator variables in cells 3 and 4 influence the form of the relationship between the predictor and criterion variables. A moderator variable that influences the form of a relationship implies a significant interaction between the moderator and predictor variables. Moderator variables in cell 3 are significantly related to the criterion or predictor variables or to both and are referred to as quasi moderators. Cell 4 moderators are not significantly related to either criterion or predictor variables and are termed pure moderators (Sharma et al., 1981).

HYPOTHESIS

Contingency theory suggests that key strategic requirements vary depending upon environmental conditions. But contingency theory provides little guidance for predicting the form or strength of relationships between strategy variables and environments. As was discussed in the preceding pages, empirical research using moderated regression analysis and subgroup analysis has generally concluded that the form of such relationships does not vary across subenvironments but that their strength does vary. On the basis of these conclusions, the following is posited:

> *Hypothesis 1: Environments moderate the strength but not the form of relationships between strategy variables and performance.*

This hypothesis states that the interaction terms of environment and strategy variables will not be statistically significant. However, the strength of the relationship between strategy variables and performance will vary across environmental subgroups. In terms of Figure 1, the hypothesis states that environment is a homologizer.

METHODS

Analysis

The analysis followed the procedures for identifying moderator variables outlined by Arnold (1982) and by Sharma and colleagues (1981). The model developed in this study identified environment as the moderator variable, a set of strategy variables as the predictor variables, and ROI as the criterion variable. An alternative model having strategy variables as moderators and environment as the predictor was rejected because, as Arnold (1982) suggested, "from the viewpoint of substantive theory, it is sensible to maintain the convention of labeling the primary causal variable under investigation as the 'independent variable' *X* and the additional variable hypothesized to moderate the causal impact of *X* as the 'moderator variable' *Z*" (1982: 154). Thus, the strategy variables were chosen as predictors because the field of strategic management focuses on managerial choices within environmental contexts. The procedure for identifying moderator variables had four steps:

Step 1: Using moderated regression analysis, determine if a significant interaction exists between the hypothesized moderator and predictor variables. If a significant interaction exists, go to step 2. Otherwise, proceed to step 3.

Step 2: Determine if the moderator variable is significantly related to the criterion variable. If it is, then the variable is a quasi moderator (cell 3, Figure 1). If not, the variable is a pure moderator (cell 4, Figure 1). The moderator variable in both cases influences the form of the relationship between the predictor and criterion.

Step 3: Determine if the moderator variable is significantly related to the predictor variable. If it is, then it is an exogeneous, predictor, intervening, antecedent, or suppressor variable (cell 1, Figure 1). If the hypothesized moderator is not related to either the predictor or criterion variables, proceed to step 4.

Step 4: Develop subgroups based on the hypothesized moderator variable. Test for the significance of predictive validity across the subgroups. If significance is found, then the variable is a homologizer operating through the error term (cell 2, Figure 1). If significant differences are not found, then the variable is not a moderator.

Data

Data for environment, strategy, and performance variables were identified for 1,638 business units in the PIMS data base over the years 1978 through 1981. Several studies have examined the quality and reliability of this data base (Anderson & Paine 1978; Hambrick, McMillan, & Day, 1982; Phillips & Buzzell, 1982; Phillips et al., 1983). The work of these authors suggests that both the quality and reliability of the data base is high.[2]

Variables

The overall measure of environment used in this study was a discrete variable with eight categories representing business units' environments. Every business unit was classified into one of eight categories according to its pattern on 17 environmental components representing characteristics of market structures. The categories were developed through cluster and discriminant analysis of the 17 environmental components (see the Appendix). The eight environmental categories are emerging, fragmented, fragmented with auxiliary services, global importing, global exporting, stable, mature, and declining. The Appendix describes their key characteristics. The eight categories are reported here because the results indicated that subgroup analysis was necessary to identify the moderating effect of the environment on the relationship between the strategy variables and performance. Further, the eight categories proved to be the most concise method of presenting the results while not distorting the analysis, results, or interpretation. The data were analyzed in three ways according to the procedure for identifying moderator variables described in the Methods section. First, each individual environmental component was analyzed separately. Second, groups of environmental components were analyzed in various combinations. Third, the eight environmental categories were analyzed together. The results were similar for all three analyses.

---

[2] See Ramanujam and Venkatraman (1984).

Strategy variables were chosen by measuring parts of strategy (Hambrick, 1980). Seventeen variables were classified into four broad theoretical dimensions of strategy, defined as follows (Hambrick, 1983): *cost efficiency*, the degree to which costs per unit of output are low; *asset parsimony*, the degree to which assets per unit of output are few; *differentiation*, the degree to which a product and its enhancements are perceived as unique; *scale scope*, the relative size and range of activities of a business within its industry.

Software limitations prohibited use of all 17 variables. Therefore, on the basis of past research using the PIMS data base and the list of strategy variables suggested by Hofer (1975) to be important for strategy research, 9 strategy variables were selected from the set of 17. Three criteria were applied: (1) representation of each of the four broad theoretical dimensions of strategy, (2) lack of multicollinearity, and (3) maximum R^2 with ROI with fewest possible variables.

Table 1 presents the set of strategy variables chosen for this study and their association with the four broad dimensions of strategy. Only 3 of the 36 intercorrelations among the strategy variables exceeded .20: capacity utilization and manufacturing expenses/revenues, −.23; employee productivity and marketing expenses/revenues, −.25; and relative product quality and relative market share, .31. Thus the set of variables demonstrated a lack of multicollinearity. In terms of maximizing R^2, the final combination of nine strategic variables accounted for 40 percent of the variance in ROI. Adding the remaining eight explained an additional 2 percent of the variance. These eight variables by themselves explained only 12 percent of the variance in ROI.

Several of the variables are arithmetically related to ROI, as is noted in Table 1. This implies that similar mathematical manipulations may partially

TABLE 1
Strategy Variables and Their Associations with Theoretical Dimensions of Strategy[a]

Strategy Variables	Associated Dimensions of Strategy
Investment/Intensity[b]	Asset parsimony
Capacity utilization	Asset parsimony
Employee productivity[b]	Cost efficiency
Relative direct cost	Cost efficiency
Manufacturing expenses/gross revenues[b]	Cost efficiency
Total R&D expenses/gross revenues[b]	Differentiation, cost efficiency, or both
Total marketing expenses/gross revenues[b]	Differentiation, cost efficiency, or both
Relative product quality	Differentiation
Relative market share	Scale/scope

[a] The four broad theoretical dimensions come from Hambrick (1983).

[b] Strategic variables with an arithmetic relationship to ROI.

account for the observed relationships between those variables and ROI (Anderson & Paine, 1978; MacMillan, Hambrick, & Day, 1982). This could potentially limit the observation of differences in the forms of the relationships between strategy variables and performance across categories of environments. In the absence of longitudinal data that would allow examination of both lagged and mathematical effects, an alternative methodological check was made. The eight remaining strategy variables from the original set were analyzed with the same procedure as our nine variables. These remaining variables were receivables/revenues, total inventory/revenues, degree of backward integration, degree of forward integration, relative compensation, percentage of purchases from the three largest suppliers, relative product line breadth, and relative price. The results of this analysis were consistent with those reported in the Results section; the Discussion section includes highlights of this analysis. Thus, although the strategy variables that were arithmetically linked to ROI may have produced R^2s that were biased upward, they do not confound the interpretation of the moderating effect of environment on the relationships between strategy variables and the performance variable, ROI.

RESULTS

This reporting of results follows the four-step procedure for identifying moderator variables outlined in the Methods section. Table 2 presents the results of a moderated regression analysis for strategy variables, environmental categories, and their interaction. The adjusted R^2s are reported. The strategy variables accounted for 40 percent of the variance in ROI; adding the environmental categories and interaction terms accounted for 2 and 0 percent respectively. Because the interaction term was not significant, the next step was to determine if environmental categories were significantly related to the strategy variables (predictors), or to ROI (criterion), or to both.

Both correlation analysis and regression analysis were used to determine if environmental categories were significantly related to ROI. The product moment correlations for each of the 17 environmental components with ROI revealed that 15 of the 17 had correlations of a magnitude less than or equal to .07 ($p > .10$). Percentage of employees unionized had a correlation of

TABLE 2
Results of Moderated Regression Analysis for Strategy Variables, Environment, and Their Interaction

Independent Variables	Cumulative R^2	F	df	ΔR^2	F	df
Strategy	.40[a]	124.0*	9,1628	.40[a]	124.00*	9,1628
Strategy and environment	.42	74.1*	16,1621	.02	5.00*	7,1621
Strategy-environment Interaction	.42	16.2*	79,1558	.00	0.00	63,1558

[a] Values reported are for adjusted R^2.

*$p < .001$

$-.11$ with ROI, and frequency of product changes had a correlation of .10. Neither of these correlations is high enough to suggest a significant relationship with ROI ($p > .10$). Regressing the eight environmental categories and the 17 environmental components separately against ROI produced adjusted R^2s of .01 ($F_{7,1630} = 2.85$, $p < .006$) and .05 ($F_{17,1620} = 5.63$, $p < .001$), respectively. We obtained separate regression equations for both the eight environmental categories and the individual environmental components to check the consistency of the results, because using aggregated categories in regression analysis often leads to a significant loss of information. The results of the two sets of regressions were consistent, indicating that the loss of information was not serious. These results, coupled with those of the moderated regressions, suggest that although the relationship between environment and ROI is statistically significant owing to large sample sizes, its practical significance is questionable.

The intercorrelations of each environmental component and strategy variable were examined to determine if they were significantly related. Of the 153 intercorrelations (17 environmental components × 9 strategy variables), four were significant ($p < .05$). These four significant correlations were industry value added and employee productivity, .42; life cycle stage and R & D expenses/revenues, $-.26$; industry exports and R & D expenses/revenues, .29; and importance of auxiliary services to end users and R & D expenses/revenues, .29. The results suggested that, overall, the environment and strategy variables were not significantly related at the $p < .05$ level.

The final step in the analysis was the identification of subgroups on the basis of the presumed moderator variable. We used the eight subenvironments referred to in the Methods section and briefly described in the Appendix for the subgroup analysis. Table 3 shows the strategy variables that had betas significantly ($p < .05$) related to ROI in each subenvironment. This table illustrates two important points. First, the sign of the betas for each strategy variable does not change across the subenvironments, which indicates a lack of interaction and is consistent with the results of the moderated regression analysis. A chi-square test (Arnold, 1982) for the strength of the correlations between each strategy variable individually and ROI across the eight subenvironments was only significant ($p < .05$) for the correlation between R & D expenses and revenues. Thus, examining the strength of the relationship between strategy variables and ROI individually across subgroups is not likely to demonstrate statistically significant differences.

The second important point that Table 3 illustrates is that the pattern of significant strategy variables varies across the eight subenvironments. A chi-square test (Arnold, 1982) for the difference in the strength of the adjusted R^2s for the set of nine strategy variables across the eight subenvironments was not significant. Therefore, the predictive validity of the set does not vary across the subenvironments. However, a Chow test (Chow, 1960) of the equality of both the slopes and the intercepts of the regression lines across the subenvironments was significant ($F = 2.08$, $df = 63, 1558$, $p < .01$). A further examination of the regression lines showed differences in their slopes

TABLE 3
Strategy Variables Significantly Related to ROI for Each Subenvironment

	Subenvironments[a]								
Strategy Variables	Exporting	Fragmented	Stable	Auxiliary Services	Emerging	Mature	Importing	Declining	Overall
Investment intensity	−[a]	−	−	−	−	−	−	−	−
Capacity utilization	+	+	+	+	+	+		+	+
Employee productivity	+		+			+		+	+
Relative direct costs	−	−	−	−		−	−	−	−
Manufacturing expenses/ revenues	−		−	−	−	−	−	−	−
Total R&D expenses/ revenues	−				−				
Marketing expenses/ revenues			−	−	−			−	−
Relative product quality	+		+	+	+	+	+		+
Relative market share			+	+		+		+	+
n	93	88	402	376	167	127	62	323	1638
R^2	.57	.46	.42	.46	.44	.52	.52	.37	.41

[a] The plus or minus represents the sign of the beta coefficient and indicates that the variable is significantly ($p < .05$) related to ROI in the sub-environment. A blank indicates that the variable is not significantly ($p < .05$) related to ROI in the subenvironment.

across the subenvironments that were at best marginally significantly different ($F = 1.77$, $df = 63, 1558$, $p < .06$). Therefore, the main difference between the regression lines was in their intercepts and not in their slopes. These results suggest that the set of strategy variables that are significantly related to ROI varies by subenvironment. This is the primary moderating effect of environments on the relationship between strategy and performance.

The results reported here support the hypothesis that environments modify the strength but not the form of the relationship between strategy variables and performance. Specifically, environment acts as a homologizer that modifies the strength of the effects of subsets of strategy variables within particular subenvironments. The subset of strategy variables that were significantly related to ROI varied across the subenvironments; however, the predictive validity of the complete set did not show statistically significant differences.

DISCUSSION

Researchers in the field of strategic management have viewed environment as a key contingency variable for strategy formulation. Thus, the finding that environment, as measured by characteristics of market structures, moderates the strength but not the form of the relationship between strategy and performance has important implications for future theory and research, public policy, and managers. One implication of this study for contingency theorists is a need to identify environmental subgroups and to examine relationships between strategy and performance within and across these subgroups. The development and identification of subgroups has become an increasingly important topic in strategy research (Hambrick, 1983, 1984; Miller & Friesen, 1984). This study developed subgroups using multiple indicators of market structures and multivariate classification techniques. Future research should explore other methods of identifying environmental dimensions, such as using perceived characteristics (Miller & Friesen) and other grouping techniques. These methods and techniques should then be examined within the framework developed by Arnold (1982) and Sharma and colleagues (1981) for their moderating influences. For example, there is some evidence that strategy process variables like managerial decision making styles moderate the forms of relationships among market share, strategy, and performance (Gupta & Govindarajan, 1984). An extension of their study would be to explore if and how environment influences these relationships.

This study also suggests that when examining parts of business-level strategy (Hambrick, 1980), researchers should examine multiple strategy variables rather than one or a few. The important finding that relationships among environment, strategy, and performance vary primarily when analysis includes multiple strategy variables supports this conclusion. Further, researchers should use multiple performance measures when possible.

The relative significance of the strategy variables within the subenvironments provides insights as to the roles of the theoretical dimensions of strategy across the subenvironments (Hambrick, 1983; see Table 1). Findings were similar to those of MacMillan and colleagues (1982) in that asset parsimony and reduced costs helped performance across all of the subenvironments. However, in the mature, declining, and global exporting subenvironments these dimensions were especially critical. The mature subenvironment further differed from the others in that relative market share, a variable associated with scope, was significant and positive while backward vertical integration, also associated with scope, was negative and significantly related to ROI. The mature subenvironment was also the only one in which relative price, associated with differentiation, was positive and significantly related to ROI. Thus, the mature subenvironment had significant strategy variables in each of the four theoretical dimensions of strategy. The stable, nonfragmented environment was similar to the mature one except that the relationships were somewhat weaker.

The fragmented subenvironment provided a situation that sharply contrasted with the mature subenvironment. As expected, the scale/scope and differentiation dimensions played small roles in the fragmented subenvironment. It appears that fragmented subenvironments lend themselves to fewer strategic options than the other subenvironments.

In the global importing subenvironment, the scale/scope dimension was found to be significant. Not only was relative market share important, but both forward and backward integration increased return on investment. This can be contrasted with results for the emerging subenvironment, in which scope was not a significant dimension, and with results for the declining subenvironment, in which scope was an important dimension but the differentiation dimension seemed to play a small role. These few observations suggest that the subenvironments can to a certain degree be contrasted and compared using the theoretical dimensions identified by Hambrick (1983). Future research should more closely examine the theoretical dimensions of strategy and how different environments influence the significance of each dimension for competitive strategy and performance.

Industrial organizational economics is extremely concerned with competition's implications for social welfare (Scherer, 1980); strategic management is more concerned with the performance of firms (Porter, 1981). The main implication of this study for public policy is that an environment's market structure should not be viewed primarily as an independent variable directly influencing the performance of firms. Rather, environment moderates the strength of relationships between business units' strategies and their performance. For example, market share, often used as a measure of market power and performance and thus as a rationale for antitrust action, may not be an appropriate indicator of either power or performance in certain contexts. This indeed was the case in this study, in which market share was significantly related to performance in only four of the eight environments. Two previous studies (Phillips et al., 1983; Prescott, Kohli, & Venkatraman, 1984) found similar results. Thus, public policy decisions need to include strategy variables for individual firms and to assess the relative effects of those variables in the environments under consideration. The finding that market structure characteristics are mildly related to ROI is consistent with results from much research in industrial organizational economics (Bothwell, Cooley, & Hall, 1984; Scherer, 1980). Scherer also suggested that economists include firm variables in their studies.

Finally, managers who develop strategies to either adapt to changing environmental conditions or to proactively influence their environments should find the results of this study encouraging. The finding that environment moderates the strength but not the form of relationships between strategy and performance certainly makes their task easier. Specifically, environment moderates the set of strategy variables significantly related to ROI. Managers should focus on identifying the strategy variables significantly related to performance in their environments and adjust their strategies accordingly. The technique of critical success factors popularized by Rockart (1979) could assist in the process. If environment does not change the form of relationships between strategy and performance, perhaps changing environmental situations require changes in the relative emphases of strategies rather than changes in strategies.

CONCLUSIONS

The relationship between the environment, strategy, and performance of a business is a central concern of strategic management. This study used moderated regression and subgroup analysis to explore the moderating influence of environment on the relationship between strategy and performance. In the results of the moderated regression analysis, a set of strategy variables accounted for 40 percent of the variance in the relationship; environment accounted for 2 percent; and the interaction term was not significant. The results of the subgroup analysis revealed that environment served as a homologizer for a set of strategy variables in that it modified the strength but not the form of the relationship between strategy and performance. Thus, environment moderates this relationship because the set of strategy variables that are significantly related to performance varies across subenvironments. These results do not suggest that environment is not important to high performance, but rather that contingency theory and research should focus on identifying meaningful subenvironments and on examining strategy-performance relationships with and across these subenvironments. Environment is critical because it establishes the context in which to evaluate the importance of various relationships between strategy and performance.

REFERENCES

Aldrich, H. 1979. ***Organizations and environments***. Englewood Cliffs, N. J.: Prentice-Hall.

Anderson, C., & Paine, S. 1978. PIMS—A reexamination. ***Academy of Management Review***, 3: 602–612.

Ansoff, H. 1980. Strategic issue management. ***Strategic Management Journal***, 1: 131–148.

Arnold, H. J. 1982. Moderator variables: A clarification of conceptual, analytic and psychometric issues. ***Organizational Behavior and Human Performance***, 29: 143–174.

Bain, J. S. 1959. ***Industrial organizations***. New York: John Wiley & Sons.

Bass, F. M. 1974. Profit and the a/s ratio. ***Journal of Advertising Research***, 14(6): 9–19.

Bass, F. M., Cattin, P., & Wittink, D. R. 1977. Market structure and industry influence on profitability. In H. B. Thorelli (Ed.), ***Strategy + structure = performance***: 181–197. Bloomington: Indiana University Press.

Bothwell, J. L., Cooley, T. F., & Hall, T. E. 1984. A new view of the market structure-performance debate. ***Journal of Industrial Economics***, 32: 397–417.

Chow, G. C. 1960. Tests of equality between sets of coefficients in two linear regressions. ***Econometrika***, 28: 591–605.

Dill, W. 1958. Environment as an influence on managerial autonomy. ***Administrative Science Quarterly***, 2: 409–443.

Everitt, B. 1974. ***Cluster analysis***. New York: John Wiley & Sons.

Gale, B. 1972. Market share and rate of return. ***Review of Economics and Statistics***, 54: 412–423.

Gupta, A. K., & Govindarajan, V. 1984. Business unit strategy, managerial characteristics, and business unit effectiveness at strategy implementation. ***Academy of Management Journal***, 27: 25–41.

Hambrick, D. C. 1980. Operationalizing the concept of business-unit strategy in research. ***Academy of Management Review***, 5: 567–576.

Hambrick, D. C. 1983. An empirical typology of mature industrial product environments. ***Academy of Management Journal***, 26: 213–230.

Hambrick, D. C. 1984. Taxonomic approaches to studying strategy: Some conceptual and methodological issues. ***Journal of Management***, 10(1): 27–42.

Hambrick, D. C. 1986. Strategies for mature industrial-product businesses: A taxonomical approach. In J. H. Grant (Ed.), ***Strategic management frontiers***: forthcoming. Greenwich, Conn.: JAI Press.

Hambrick, D. C., MacMillan, I. G., & Day, D. L. 1982. Strategic attributes and performance in the BDG matrix—A PIMS-based analysis of industrial product businesses. ***Academy of Management Journal***, 25: 510–531.

Hatten, K. J., & Schendel, D. E. 1977. Heterogeneity within an industry: Firm conduct in the U. S. brewing industry, 1952–1971. ***Journal of Industrial Economics***, 26: 97–113.

Hatten, K. J., Schendel, D. E., & Cooper, A. C. 1978. A strategic model of the U. S. brewing industry: 1952–1971. ***Academy of Management Journal***, 21: 592–610.

Hitt, M. A., Ireland, R. D., & Stadter, G. 1982. Functional importance and company performance: Moderating effects of grand strategy and industry type. ***Strategic Management Journal***, 3: 315–330.

Hofer, C. W. 1975. Toward a contingency theory of business strategy. ***Academy of Management Journal***, 18: 784–810.

Hofer, C. W., & Schendel, D. 1978. ***Strategy formulation: Analytical concepts***. New York: West Publishing Co.

Jauch, L. R., Osborn, R. W., & Glueck, W. F. 1980. Short-term financial success in large business organizations: The environment-strategy connection. ***Strategic Management Journal***, 1: 49–63.

Lawrence, P., & Lorsch, J. 1967. ***Organization and environment***. Boston: Harvard University Press.

Lazarsfeld, P. F. 1955. Interpretations of statistical relations as a research operation. In P. F. Lazarsfeld & M. Rosenberg (Eds.), ***The language of social research***: 115–124. Glencoe, Ill.: The Free Press.

Lenz, R. T., & Engledow, J. L. 1986. Environmental analysis: The applicability of current theory. ***Strategic Management Journal***, 7(1): in press.

MacMillan, I. C., Hambrick, D. C., & Day, D. L. 1982. The product portfolio and profitability—A PIMS-based analysis of industrial-product businesses. ***Academy of Management Journal***, 25: 733–755.

Miles, R. E., & Snow, C. C. 1978. ***Organizational strategy, structure, and process***. New York: McGraw-Hill Book Co.

Miller, D., & Friesen, P. H. 1984. ***Organizations: A quantum view***. Englewood Cliffs, N. J.: Prentice-Hall.

Naisbitt, J. 1982. ***Megatrends***. New York: Warner Books.

Neter, J., & Wasserman, W. 1974. ***Applied linear statistical models***. Homewood, Ill.: Irwin.

Newman, H. H. 1978. Strategic groups and the structure-performance relationship. ***Review of Economics and Statistics***, 60: 417–427.

Oster, S. 1982. Intra-industry structure and the ease of strategic change. ***Review of Economics and Statistics***, 63: 376–383.

Pfeffer, J., & Salancik, G. R. 1978. ***The external control of organizations***. New York: Harper & Row Publishers.

Phillips, L. W., & Buzzell, R. D. 1982. ***An examination of the reliability of PIM'S competitive strategy measures***. Working paper, Stanford University, Palo Alto, Calif.

Phillips, L. W., Chang, D. R., & Buzzell, R. D. 1983. Product quality, cost position and business performance: A test of some key hypotheses. ***Journal of Marketing***, 47(2): 26–43.

Porter, M. E. 1979. The structure within industries' and companies' performance. ***Review of Economics and Statistics***, 61: 214–227.

Porter, M. E. 1980. ***Competitive strategy***. New York: The Free Press.

Porter, M. E. 1981. The contributions of industrial organization to strategic management. ***Academy of Management Review***, 6: 609–620.

Prescott, J. E. 1983. ***Competitive environments, strategic types, and business performance: An empirical analysis***. Unpublished doctoral dissertation, Pennsylvania State University, University Park, Pa.

Prescott, J. E., Kohli, A. J., & Venkatraman, N. 1984. Is the relationship between market share and business profitability spurious?: An empirical assessment. ***Academy of Management Proceedings***: 32–36.

Ramanujam, V., & Venkatraman, N. 1984. An inventory and critique of strategy research using the PIMS data base. ***Academy of Management Review***, 9: 138–151.

Rockart, J. F. 1979. Chief executives define their own data needs. ***Harvard Business Review***, 57(2): 81–93.

Rosenberg, M. 1968. ***The logic of survey analysis***. New York: Basic Books.

Scherer, F. M. 1980. ***Industrial market structure and economic performance***. Chicago: Rand McNally & Co.

Schoonhoven, C. 1981. Problems with contingency theory: Testing assumptions hidden within the language of contingency. ***Administrative Science Quarterly***, 26: 349–377.

Sharma, S., Durand, R. M., & Gur-Arie, O. 1981. Identification and analysis of moderator variables. ***Journal of Marketing Research***, 18: 291–300.

Snow, C. C., & Miles, R. E. 1983. The role of strategy in the development of a general theory of organizations. In R. Lamb (Ed.), ***Advances in strategic management***, vol. 2: 231–259. Greenwich, Conn.: JAI Press.

Weick, K. 1979. ***The social psychology of organizing*** (2nd ed.). Reading, Mass.: Addison-Wesley.

Wishart, D. 1978. ***Clustan*: *User manual***. Edinburgh, Scotland: Program Library Unit, Edinburgh University.

Zedeck, S. 1971. Problems with the use of moderator variables. ***Psychological Bulletin***, 76: 295–310.

APPENDIX

The empirical development of eight environmental categories or subgroups based on 17 characteristics of market structures employed cluster analysis; results were interpreted in the context of Porter's (1980) theoretical set of generic industries. The first five steps identified eight relatively independent and homogeneous subgroups. The final two steps involved interpreting the subgroups.

Step 1 was selection of 17 environmental components based on theory and previous research (Scherer, 1980) and lack of multicollinearity.

Step 2 was random selection of 311 business units.

Step 3 was cluster analysis of the 311 business units (Ward's method; see Hambrick, 1983, and Wishart, 1978).

Step 4 was chosing the number of subgroups. The criteria were: (1) examination of sharp changes in error sum of squares when the number of clusters is changed, and (2) visual inspection of the dendogram (Everitt, 1974). Eight groups were identified.

Step 5 was cross validation. Discriminant functions were developed for each of the eight groups identified in step 4. The discriminant functions were then used to classify 1,792 business units into the eight subenvironments. For a business unit to remain in the sample, it had to have a probability of group membership in any one group of .5 or greater. This criterion was chosen to ensure homogeneous groups. Of the 1,792 business units, 1,638 (91%) met the criterion. The average probability of group membership for the 1,638 business units was 88 percent.

Step 6 was developing profiles for each subgroup based on both natural and standardized mean scores of the 17 environmental components. Descriptive statistical data on these components are available from the author.

Step 7 was interpreting the subgroups in terms of Porter's (1980) typology of generic industries.

The following describes key characteristics for the eight subgroups. Descriptive statistical data are available from the author.

1. Global exporting
   ROI = 23.88%. Key characteristics include: greater than 30 percent of sales represent exports; growth stage of product life cycle; 12 percent long-term industry growth; 10 percent real market growth; and auxiliary services important to end users.
2. Fragmented (standard products)
   ROI = 14.51%. Key characteristics include: fragmented set of customers; annual product changes; purchases by end users annually; and low industry concentration.
3. Stable, nonfragmented
   ROI = 23.01%. Key characteristics include: high market share stability among competitors; very little customer fragmentation as evidenced by 10 percent of customers accounting for 50 percent of sales; purchase frequency is once per month; and new product development less than two years.
4. Fragmented with auxiliary services
   ROI = 22.28%. Key characteristics include: low industry concentration; fragmented customers; purchases by end users annually; and high importance of auxiliary services.

5. Emerging
   ROI = 20.31%. Key characteristics include: high market share instability; early in product life cycle; 13 percent long-term industry growth; 19 percent real market growth; low percentage of employees unionized; and no regular pattern of product changes.
6. Mature
   ROI = 23.32%. Key characteristics include: industry concentration around 70 percent; mature stage of product life cycle; high growth rates for both materials and wages; and high percentage of unionized workers.
7. Global importing
   ROI = 28.69%. Key characteristics include: greater than 30 percent of sales represent imports; high concentration is 74 percent; and long term growth is 12 percent.
8. Declining
   ROI = 20.39%. Key characteristics include: decline stage of product life cycle; real market growth is less than 1 percent; high value added/employee; high purchase frequency; long development time for new products; and auxiliary services unimportant.

John E. Prescott earned his Ph. D. degree at the Pennsylvania State University; he is an assistant professor of business administration at the Graduate School of Business, University of Pittsburgh. His current research interests include the interfacing of industrial organizational economics and strategy management, and industry and competitor analysis.

© Academy of Management Journal
1986, Vol. 29, No. 2, 347–372.

MULTI-OBJECTIVE ASSESSMENT OF EFFECTIVENESS OF STRATEGIC PLANNING: A DISCRIMINANT ANALYSIS APPROACH

VASUDEVAN RAMANUJAM
Case Western Reserve University
N. VENKATRAMAN
Massachusetts Institute of Technology
JOHN C. CAMILLUS
University of Pittsburgh

Drawing on the relevant literature, this study develops seven key dimensions of planning systems, five reflecting their design aspects and two tapping the organizational context of planning. Discriminant analysis is employed to evaluate the ability of these seven dimensions to distinguish between more and less effective planning systems, using three different criteria of planning effectiveness. Although influences on effectiveness seem to vary depending on the criterion of effectiveness used, overall the most important dimensions were system capability, or a system's orientation toward creativity and control; resources provided for planning; and functional coverage.

The relationship between strategic planning, which we will refer to as simply planning in this study, and organizational performance is one of the most extensively researched issues in the field of strategic management. An illustrative list of studies addressing the effects of planning on performance includes Ansoff, Avner, Brandenburg, Portner, and Radosevich (1970), Fulmer and Rue (1974), Grinyer and Norburn (1975), Herold (1972), Karger and Malik (1975), Kudla (1980), Lindsay and Rue (1980), Malik and Karger (1975), Rue and Fulmer (1973), Thune and House (1970), and Wood and LaForge (1979). Two studies provide complete reviews and critiques of this literature (Armstrong, 1982; Lorange, 1979).

Despite the large number of research efforts that attempt to elucidate links between planning and performance, the results of this body of research are fragmented and contradictory. This state of affairs has arisen because most previous research on the effectiveness of planning systems has suffered

---

Portions of this paper were presented at the Fourth Annual Strategic Management Society Conference, Philadelphia, in October 1984. The authors would like to thank Jarmell Boyd, Patrick Gaughan, Linda Natal, and Loretta Riles for their administrative assistance during the data collection and analysis stages of the project on which this paper is based. The insightful comments and suggestions provided by the two anonymous reviewers are also gratefully acknowledged.

from two major conceptual shortcomings. First, most studies have used rather simplistic conceptualizations of the notion of planning. Thus, researchers have attempted to show differences in financial performance between "planners" and "nonplanners" or "formal planners" and "informal planners." Second, most studies have been solely preoccupied with the linkage between planning and the financial aspects of corporate performance, even though conceptual writings on formal planning systems stress several nonfinancial, intangible benefits (Camillus, 1975; Steiner, 1979).

With regard to the first issue, crude dichotomizations along single dimensions like the presence or absence of planning or its degree of formality are clearly suspect, because the debate as to what the term planning connotes is still a live one (Mintzberg, 1981; Snyder, 1982). Such conceptualizations are inconsistent with the multidimensional view of planning systems that is becoming more and more visible in the literature (e.g., King & Cleland, 1978; Lorange, 1979, 1980; Lorange & Vancil, 1977). Similarly, with regard to the second issue, an exclusive emphasis on financial performance is conceptually flawed and is open to attack as being narrowly focused. A broader concept of effectiveness is needed.

A truly meaningful assessment of the value of planning systems should, therefore, recognize their multidimensional nature and the plurality of approaches that can be used to assess their worth. This research study attempted such an assessment; it has three distinctive features: First, on the basis of an extensive literature review, we identify seven distinct design and contextual dimensions of planning systems that many previous researchers believe influence their effectiveness. Second, we explicitly recognize the many possible meanings of effectiveness in the context of planning and use multiple criteria for assessing the effectiveness of planning systems. Third, we cast the study in a comparative mode by using the technique of discriminant analysis in order to bring the differences between more effective and less effective planning systems into sharp relief.

The purpose of this paper is to test a two fold proposition: (1) relatively more effective systems will differ from relatively less effective ones along the seven key design and contextual dimensions referred to earlier in this section, and (2) the relative importance of these dimensions will vary depending upon the criterion of effectiveness used.

DIMENSIONS OF PLANNING SYSTEMS

Planning systems are multifaceted management systems that are contextually embedded. Hence, they cannot be adequately described in terms of one or two characteristics such as "formality." Although employing different terminology, most writings (e.g., King & Cleland, 1978; Lorange & Vancil, 1977; Steiner, 1979) have emphasized the notion that planning systems are best described in multidimensional terms. Unfortunately, there is at yet no consensus as to what these dimensions are.

Nevertheless, in one way or other, most researchers emphasize two sets of influences on a planning system's effectiveness. These are the design elements of the planning system and the organizational context of planning. We synthesized seven dimensions that tap these two aspects. Five design elements were included in this study. They are (1) system capability, (2) use of techniques, (3) degree of attention to internal facets, (4) degree of attention to external facets, and (5) functional coverage. The following two dimensions captured the organizational context of planning: (1) resources provided for strategic planning, including top management support, and (2) resistance to planning or "anti-planning biases" (Steiner, 1979: 95–108). Although we could identify other dimensions that might conceivably influence the effectiveness of planning, these seven have the support of an extensive literature. Also, they are variables that are relatively amenable to control by managers.

Table 1 summarizes the seven dimensions and indicates the support in the literature for each of them. The following paragraphs provide a brief discussion of each dimension.

Design Elements

System capability. Strategic planning systems differ in the extent to which they emphasize creativity as opposed to control (Camillus, 1975; Shank, Niblock, & Sandalls, 1973). It has been stressed that these two somewhat opposed orientations need to be balanced in order to enhance the effectiveness of a system. An excessive emphasis on creativity at the expense of control or vice versa is apt to be dysfunctional. We conceptualize system capability in terms of the ability of a system to foster control as well as creativity. The creative capability of a system is assessed in terms of its ability to anticipate surprises and crises, its ability to adapt to unforeseen changes, and so forth. On the other hand, the control aspect of a system refers to the degree of emphasis given to managerial motivation, upward and downward communication within an organizational hierarchy, integration of diverse operational areas, and the like.

Use of techniques. A variety of analytical techniques and methodologies have been developed to aid managers in identifying and dealing with strategic decisions and problems (Grant & King, 1979). Use of these techniques is one indication of the extent of formalization of a planning process. The extent of reliance on various planning techniques and methodologies also indicates what an organization's approach to decision making is; for instance, a comprehensive versus an incremental approach might be indicated (Fredrickson, 1984).

The extent of reliance on planning techniques is thus an important dimension of planning systems. Since organizations use formal techniques with the assumption that these techniques will enhance effective definition and resolution of the ill-structured problems of planning, we also expect that over time this dimension will influence the effectiveness of a planning system itself.

Attention to internal facets. A formal approach to planning usually begins with an assessment of an organization's recent history and current situation.

TABLE 1
Dimensions of Planning Systems

Dimensions	Description	Key Supporting Literature
Design elements		
System capability	The ability of a formal planning system to balance creativity and control; adaptive flexibility of a system and its capability to support strategy formulation and implementation.	Ansoff (1975, 1984) Anthony & Dearden (1976) Camillus (1975) Lorange & Vancil (1977) King & Cleland (1978) Thompson (1967)
Use of techniques	The degree of emphasis given to the use of planning techniques to structure ill-defined, messy, strategic problems.	Grant & King (1979, 1982) Hofer & Schendel (1978) Hax & Majluf (1984)
Attention to internal facets	The degree of attention to internal (organizational) factors, past performance, and analysis of strengths and weaknesses.	Camillus & Venkatraman Grant & King (1982) King & Cleland (1978) Lorange & Vancil (1977) Stevenson (1976)
Attention to external facets	The level of emphasis given to monitoring environmental trends..	Aguilar (1965) Fahey & King (1977) Keegan (1974) Kefalas & Schoderbek (1973) Thomas (1980)
Functional coverage	The extent of coverage given to different functional areas with a view to integrating different functional requirements into a general management perspective.	Hitt, Irland, & Palia (1982) Hitt, Ireland, & Stadter (1982) Lorange (1980) Snow & Hrebiniak (1980)
Organizational context of planning		
Resources provided for planning	The degree of organizational support in the form of number of planners, involvement of top management in planning, etc.	King & Cleland (1978) Steiner (1979)
Resistance to planning	The need to anticipate and overcome resistance to planning and to create a favorable climate for effective planning.	Steiner (1979) Steiner & Schollhammer (1975) Schultz & Slevin (1976)

Previous studies have variously referred to this step or stage in planning as a "situation audit," an "appraisal," and so forth (Ansoff, 1965; Steiner, 1979). In this assessment, organizations' internal capabilities receive close study to identify their strengths and weaknesses. Often plans fail because of an inadequate or incorrect assessment of an organization's internal aspects (King & Cleland, 1978). This study treats degree of attention to internal aspects such as past performance, current strengths and weaknesses, and diagnostic assessment of recent failures or performance shortfalls as a distinct and important influence on the effectiveness of an organization's planning.

Attention to external facets. One of the purposes of planning is to adapt organizations to environmental demands and pressures. Analyzing external opportunities and threats is a major element of planning (Andrews, 1971). There is no need to belabor the importance the literature grants external orientation as an influence on the effectiveness of planning. This study therefore includes the extent of attention devoted to various aspects of environments as a further aspect of the design of planning systems.

Functional coverage. This term refers to the degree of emphasis accorded to various functional areas in planning. Functional coverage can vary because of strategic differences in the competitive postures of firms in an industry. Some firms may attempt to compete on the basis of price and volume, and others may emphasize product differentiation and customer service. The distinctive competencies demanded by these alternate approaches are usually manifested as patterns of emphasis on different functional areas (Hitt, Ireland, & Palia, 1982; Snow & Hrebiniak, 1980). In the first case, manufacturing efficiencies may be key, but in the second case the marketing function may receive special emphasis. A balanced emphasis on all functions may be more important in yet other situations in which general management is the key success factor. Therefore, we included the degree of emphasis given to various functional areas as a dimension of the design of planning systems.

Organizational Context of Planning

Resources provided for planning. Many authors have emphasized that planning in an organization cannot be successful unless adequate resources are committed to that activity (King & Cleland, 1978; Steiner, 1979). These resources may be physical—a separate planning staff and office; or they may be intangible—the time spent by the CEO and other key managers in the planning function. Planning is not a costless activity. If organizations expect benefits from planning, they must also be willing to incur commensurate levels of tangible and intangible costs of doing effective planning.

Resistance to planning. Early research on planning systems emphasized the importance of identifying and overcoming sources of resistance to organizational planning (Steiner, 1979; Steiner & Schollhammer, 1975). Organizational members may show resistance to a system in many ways; some examples are lack of acceptance of the outputs of planning, withdrawal or nominal participation in planning activities without active involvement on the part of operating managers, and gaming behaviors. Resistance to the idea and processes of planning can be expected to exert a negative influence on the effectiveness of planning systems.

MULTIPLE CRITERIA OF EFFECTIVENESS

This study approached the effectiveness of planning systems from three perspectives. First, we examined the extent of fulfillment of key planning objectives. Second, in keeping with most prior research, we included an organization's economic performance as a further criterion of effectiveness. Third, we applied an overall measure of satisfaction with a planning system.

Fulfillment of Planning Objectives

Six commonly emphasized objectives of planning were considered. These were (1) predicting future trends, (2) evaluating alternatives, (3) enhancing management development, (4) avoiding problem areas, (5) improving short-term performance, and (6) improving long-term performance. The following subsections discuss and justify each of these objectives.

Predicting future trends. Planning is often justified on the grounds that contemporary organizational environments have become increasingly turbulent, thereby necessitating some formal mechanisms for environmental monitoring and coping with environmental change (Ansoff, 1984; Gluck, Kaufman, & Walleck, 1980, 1982). An objective that follows from this reasoning is improvement of the ability of a planning system to predict future trends. Proponents of such capabilities do not suggest that planning should necessarily result in accurate predictions of the future, but that planning should at least help organizations to delineate probable, plausible, and preferable future states of the world (Amara, 1981). In fact, according to one investigation (Paul, Donavan, & Taylor, 1978), one of the major problems with planning is the inability of planners to produce reasonably valid forecasts of the future. Thus, predicting future trends is considered an important task of planning and constitutes our first objective.

Evaluating alternatives. Critics of planning always point out the tendency of planning to degenerate into a numbers game. A good planning system, it has been proposed, should not only serve as a vehicle for mind-stretching (Camillus, 1975), but also should delicately balance control and creativity (Shank, Niblock, & Sandalls, 1973). The extent to which a planning system fosters the creative generation and exploration of alternative courses of action thus becomes a further important test of its effectiveness, and it is second in our list of objectives.

Avoiding problem areas. Yet another approach to the evaluation of planning systems is to examine the extent to which they result in an accumulation of experience and enhance organizational learning (Shrivastava & Grant, 1985). An evolutionary approach to systems design (Lorange & Vancil, 1977) is consistent with the idea that effective planning systems are adaptive learning systems. Learning can be said to be occurring if a system increases the probability of achieving goals and minimizes recurrence of errors. A corollary is that an effective planning system should enable an organization to avoid problem areas. This is the third objective for planning.

Enhancing management development. Formalization of planning should lead to an improvement in the quality of management and facilitate management succession. This constitutes the long-term educational value of the planning process (Hax & Majluf, 1984; Lorange & Vancil, 1977). Planning's contribution to the development of the depth of management in an organization is by no means an easy achievement to evaluate. Yet, some attention to this important objective appears to be warranted in evaluating planning systems. This is the fourth objective.

Improving short-term and long-term performance. Finally, the importance of performance, both short-term and long-term, cannot be overemphasized. In a sense, improving performance is the *raison d'etre* for adopting elaborate planning systems in the first place. This is implicit in the exclusive focus on performance that characterizes the impressive body of studies on evaluation of planning systems identified earlier. These are the fifth and the sixth objectives.

Performance Relative to Competition

In addition to the achievement of key planning objectives, we posited that effective systems would improve organizational performance. There is a subtle difference between an organization's achieving performance objectives and performing at a relatively higher level. Hence, we included relative performance as a separate criterion of planning effectiveness. This study used four performance indicators, each measured relative to competitors rather than in absolute terms. These indicators are growth in sales, growth in earnings, changes in market share, and return on investment.

Satisfaction with Planning Systems

Adopting the approach common in the literature concerning the implementation of management information systems (Lucas, 1978), we treated satisfaction with planning systems as an additional criterion of effectiveness. Satisfaction is particularly important when use of a system is mandatory rather than voluntary, which is likely to be the case for an organization's planning system. Satisfaction with a system is likely to reflect fulfillment of objectives; thus, satisfaction can also be regarded as an additional internal validation criterion for the variables measuring fulfillment of objectives.

RESEARCH METHODS

Data Collection

A detailed questionnaire on their companies' planning practices was mailed to 600 executives of *Fortune* 500, *Fortune* 500 Service, and *Inc* 500 companies. We received completed questionnaires from 207 executives representing both planning and other operating functions, for a response rate of 34.5 percent. This level of response from a list of predominantly *Fortune* 1000 firms is higher than is usual for mail surveys of this population (Gaedeke & Tootelian, 1976). Table 2 shows characteristics of respondents and their companies.

The data show bias in favor of larger firms, with 60.9 percent of responding firms having annual sales in excess of $1 billion. Similarly, planning executives (68.5%) and respondents at the corporate levels of their organizations (81.5%) predominated.

TABLE 2
Characteristics of Respondents and Their Companies[a]

Characteristics	All Respondents (n = 207)	Responses Used in This Study[b] (n = 93)
Sales of responding units		
Less than $250 million	17.4	24.4
$251 million - $1 billion	21.8	21.1
Over $1 billion	60.9	54.4
Primary business categories of responding units		
Consumer goods	19.2	13.0
Capital goods	18.7	20.7
Raw or semifinished materials	12.8	10.9
Components for finished goods	8.9	12.0
Service sector	40.4	43.5
Organizational level of responding units		
Corporate	81.5	80.4
Business unit	18.5	19.6
Title/job responsibilities of responding executives		
Planning	68.5	67.0
Other line function	31.5	33.0
Maturity of planning systems (Year planning began)		
Before 1960	6.2	5.9
1961-1975	36.7	37.6
After 1976	55.2	56.5

[a] All figures are percentages. Nonrespondents have been excluded.

[b] A subset of respondents was used for analysis in this study because of missing data. On the basis of chi-square tests, the subset used is not statistically different from all respondents with respect to any of the five company and respondent characteristics summarized in this table, except sales categories.

Since smaller firms do not usually undertake planning to any significant extent (Robinson & Pearce, 1984), this bias toward largeness is not likely to be a serious detriment to drawing tentative generalizations about the planning processes of firms that do plan. A comparison of the responses of operating executives and planning executives using t-tests indicated that planning executives as a group did not tend to systematically overrate or underrate their planning systems. However, since a single respondent from each organization provided data, there is a real possibility of response bias or selection bias. A trade-off between number of respondents overall and number of responses per firm became inevitable in view of limitations of our resources, and we leaned in favor of a large overall number of respondents. Because our target respondents were top-level planning or operating executives, we assumed that they would be quite knowledgeable about their planning systems and generally unbiased in their responses.

Measurements

The seven systems dimensions were measured by multi-item scales. The lack of valid and reliable measurement scales has been one of the impediments to the development of middle-range theories of strategy and strategic planning (Venkatraman & Grant, 1986). This study sought to assure content validity of each dimension by use of multiple experts, including the authors, and an iterative procedure for ensuring exhaustive coverage of each construct's domain. Our use of multi-item scales was motivated by the aim of enhancing the reliability of measurements (Nunnally, 1978).

Effectiveness was measured using both single-item and multi-item scales. The six individual variables measuring fulfillment of objectives and their simple additive composite made up the objective fulfillment battery. The four relative performance items and their composite constituted the relative performance battery. Satisfaction was measured by a single-item scale. All variables were measured on appropriately anchored 5-point Likert scales.

The Appendix describes the multi-item scales used in this study. Table 3 reports means, standard deviations, reliabilities, and intercorrelations for the planning system dimensions. Table 4 shows corresponding statistics for the variables measuring effectiveness and composites.

It should be noted that the dimensions of planning systems used in this study were specified a priori because of extensive support from the literature (see Table 1). Nevertheless, we undertook factor analysis of the variable set to determine if any further parsimony could be accomplished in the number of dimensions used. We prepared analyses ranging from a three- to a seven-factor solution and carefully studied the varimax-rotated factor scores. Since

TABLE 3
Means, Standard Deviations, Reliabilities, and Intercorrelations of the Seven Dimensions of Planning Systems[a]

Dimensions	Means	s.d.	1	2	3	4	5	6	7
1. System capability	3.65	0.62	(93)						
2. Use of techniques	3.18	0.45	60	(71)					
3. Attention to internal facets	3.36	0.61	52	45	(68)				
4. Attention to external facets	3.62	0.50	45	22	19	(51)			
5. Functional coverage	3.57	0.36	39	28	34	47	(36)[b]		
6. Resources provided for planning	3.61	0.74	64	49	36	36	40	(76)	
7. Resistance to planning	1.66	0.61	−69	−43	−28	−41	−33	−70	(70)

[a] All values are based on data from 93 firms used in the discriminant analyses. Decimal points are omitted for reliabilities and correlations. Reliabilities (Cronbach alphas) are shown within parentheses. All correlations are significant at $p < .05$.

[b] The reliability for this dimension for all respondents ($n = 207$) was 0.77. Although there were differences for the other dimensions as well, they were less pronounced.

TABLE 4
Means, Standard Deviations, and Intercorrelations of the Variables Measuring Effectiveness of Planning Systems[a]

Variables	n	Means	s.d.	1	2	3	4	5	6	7	8	9	10	11	12	13
1. Predicting future trends	93	3.33	0.86	–												
2. Evaluating alternatives	93	3.75	0.72	35	–											
3. Avoiding problem areas	93	3.45	0.83	41	59	–										
4. Enhancing management development	92	3.26	0.91	69	43	41	–									
5. Improving short-term performance	93	3.46	0.84	38	34	26	37	–								
6. Improving long-term performance	92	3.67	0.76	40	54	41	50	31	–							
7. Objective composite[b]	91	3.48	0.59	77	72	70	80	63	72	–						
8. Growth in sales	87	3.61	0.99	29	11	14	26	32	23	35	–					
9. Growth in earnings	86	3.58	1.18	14	01	13	12	41	13	25	60	–				
10. Changes in market share	86	3.49	0.95	26	09	04	26	36	25	32	69	54	–			
11. Return on investment	85	3.40	1.17	15	−06	03	15	35	09	19	52	79	51	–		
12. Performance composite[c]	85	3.51	0.90	20	03	09	22	45	19	32	82	89	79	86	–	
13. Satisfaction with the planning system	93	3.41	0.99	28	37	32	32	38	29	45	31	33	15	21	29	–

[a] All correlations (decimal points are omitted) above r = 0.18 are significant at $p < .05$.

[b] The reliability of this composite was 0.82.

[c] The reliability of this composite was 0.86.

none of the factor structures were easy to interpret, we decided to rely on the seven originally specified dimensions.

Choice of Analytic Technique

Given our interest in exploring the dimensions of systems contributing to differences in effectiveness between more and less effective systems, we deemed discriminant analysis to be the appropriate analytical approach. Although use of another statistical technique, such as multiple regression, may seem more appropriate because the effectiveness variables are interval measures, an examination of the correlation matrix for the set of independent variables revealed significant multicollinearity in the data. This further reinforced our original decision to use discriminant analysis as the analytical method, because multicollinearity does not affect interpretation of the results of discriminant analysis, despite what many researchers believe (Eisenbeis, 1977).[1]

Groupings for discriminant analysis. The effectiveness groupings for discriminant analysis were constructed as follows. Since responses were made on 5-point scales, all respondents choosing 1, 2, or 3 were placed in one group, and those checking 4 or 5 were placed in the other group.[2] In the case of the two composites, we used the median as the cut point to yield the two groups for discriminant analysis.

Tests of multivariate normality. The accuracy of the classifications that result from applying discriminant analysis are sensitive to the assumption that the predictor variables constitute a multivariate normal population. To test for this, we employed the z-tests for the significance of the coefficients of skewness and kurtosis discussed by Bock (1975). The discriminating variables used in this study displayed significant deviation from normality,[3] suggesting caution in interpreting the accuracy of the classifications reported. However, our primary concern was with descriptive rather than predictive relevance (Frank, Massy, & Morrison, 1965). We therefore do not believe that the failure of a key model assumption, which would be critical for prescriptive studies, constitutes a serious challenge.

Analysis. A total of 13 discriminant analysis runs were performed. Groupings in seven runs were based on the variables measuring fulfillment of objectives and their composite; in five runs, the performance variables and their composite; and in one run, the measure of overall satisfaction. As seven discriminators were composite dimensions, missing data on even one of the component variables for any respondent led to the loss of that case from our analysis; 93 cases had complete data for constructing all the seven planning system dimensions. A comparison of these 93 cases with the 114 incomplete cases with respect to various characteristics confirmed that this data reduction did not lead to any significant biasing in the subset used for discriminant analysis (Table 2).

---

[1] Thirteen regressions were nevertheless obtained. They yielded statistically significant models with multiple R values in the range 0.42 to 0.75. However, the signs of the coefficients of the independent variables were uninterpretable because of multicollinearity in the data, which is likely due to common method variance arising from our method of data collection.

[2] Before assigning the neutral responses (scored 3) to the low effectiveness groups, we ran ANOVAs for each grouping variable for all seven discriminating dimensions. Scheffé pairwise comparisons across the three groupings showed that the neutral groups were more similar to the low effectiveness groups than to the high effectiveness groups, thus justifying the pooling of the neutral ratings with the low effectiveness ones.

[3] This problem persisted even when various data transformations were tried. The detailed results of this analysis are not reported here in the interest of conserving space.

RESULTS AND DISCUSSION

Overall Results

Table 5 shows the results for the discriminant analysis using satisfaction as the effectiveness criterion and serves as a prototype for Tables 6 and 7, which summarize the results for the variables measuring fulfillment of objectives and relative performance, respectively.

Of the 13 discriminant functions obtained, 12 were significant at $p < .05$. The run for growth in earnings was significant at $p < .10$. It is encouraging that the same seven dimensions lead to significant discriminant functions for each of the effectiveness criteria used. To a degree, this is not surprising. The variables measuring fulfillment of objectives were all highly intercorrelated, as were those measuring relative performance (Table 4). In addition, the objective fulfillment variables had high correlations with the

TABLE 5
Results of Discriminant Analysis for Groupings Based on Satisfaction

Criteria	Results
Group sizes	
Number of dissatisfied planners (group 1)	42
Number of satisfied planners (group 2)	51
Significance level of the linear discriminant function	.01
Assumption of equality of group dispersion matrices (p for Box's M)	Not valid (.01)
Percent classified accurately by linear classification rule[a]	
Group 1	78.6
Group 2	84.3
Overall	81.7
Percent accuracy of chance model based on sample group prior probabilities	50.5
Standardized discriminant function coefficients	
System capability	−0.33
Use of techniques	−0.07
Attention to internal facets	0.05
Attention to external facets	0.50
Functional coverage	−0.17
Resources provided for planning	−0.23
Resistance to planning	0.71

[a] The failure of Box's *M* test for equality of group dispersion matrices suggests that quadratic classification procedures might be more appropriate. The use of a quadratic clarification rule, however, did not lead to a noticeable improvement in classification accuracies.

TABLE 6
Results of Discriminant Analysis for Groupings Based on Variables Measuring Fulfillment of Objectives

	Measures of Fulfillment of Objectives						
Criteria	Predicting Future Trends	Evaluating Alternatives	Avoiding Problem Areas	Enhancing Management Development	Improving Short-term Performance	Improving Long-term Performance	Objective Composite
Group Sizes							
Group 1: entirely unfulfilled, somewhat unfulfilled, or neutral	42	22	43	48	46	30	46
Group 2: somewhat fulfilled or entirely fulfilled	51	71	50	44	47	62	45
Significance levels of linear discriminant functions	.01	.05	.05	.01	.01	.01	.01
Assumption of equality of group dispersion matrices (p for Box's M)	Not valid (.05)	Not valid (.01)	Valid (ns)	Valid (ns)	Valid (ns)	Valid (ns)	Not Valid (.05)
Percent classified accurately by linear classification rule[a]							
Group 1	64.3	68.2	69.8	81.2	65.2	66.7	73.9
Group 2	68.6	67.6	62.0	72.7	72.3	74.2	71.1
Overall	66.7	67.7	65.6	77.2	68.8	71.7	72.5

Criteria	Measures of Fulfillment of Objectives						
	Predicting Future Trends	Evaluating Alternatives	Avoiding Problem Areas	Enhancing Management Development	Improving Short-term Performance	Improving Long-term Performance	Objective Composite
Percentage accuracy of chance model based on sample group prior probabilities	50.5	63.9	50.3	50.1	50.0	56.0	50.0
Standardized discriminant function coefficients							
System capability	0.51	−0.90	1.17	0.65	0.00	0.08	0.35
Use of techniques	0.59	0.05	−0.56	−0.07	0.82	−0.42	0.27
Attention to internal facets	−0.16	0.21	0.33	0.27	−0.10	0.14	0.27
Attention to external facets	−0.51	0.54	−0.51	−0.12	0.21	0.38	−0.22
Functional coverage	0.42	−0.02	0.16	0.49	0.09	0.22	0.34
Resources provided for planning	−0.11	−0.15	−0.31	−0.39	0.69	0.46	−0.03
Resistance to planning	−0.21	0.35	−0.19	−0.40	0.77	−0.35	−0.36

[a] The failure of Box's *M* test for equality of group dispersion matrices suggests that quadratic classification procedures might be more appropriate in three of the analyses. The use of quadratic classication did not, however, lead to a noticeable improvement in classification accuracies.

TABLE 7
Results of Discriminant Analysis for Groupings Based on Performance Relative to Competition

	Performance Measures				
Criteria	Growth in Sales	Growth in Earnings	Changes in Market Share	Return on Investment	Performance Composite
Group sizes					
Group 1: Equal, worse, or much worse	36	33	42	40	43
Group 2: Better or much better	51	53	44	45	42
Significance level of the linear discriminant function	.05	.10	.01	.01	.01
Assumption of equality of group dispersion matrices (p for Box's M)	Not valid (.01)	Valid (.10)	Not valid (.01)	Valid (ns)	Not valid (.01)
Percent classified accurately by linear classification rule[a]					
Group 1	69.4	63.6	71.4	70.0	72.1
Group 2	68.6	60.4	72.7	75.6	71.4
Overall	69.0	61.6	72.1	72.9	71.8
Percentage accuracy of chance model based on sample group prior probabilities	51.5	52.7	50.0	50.2	50.0
Standardized discriminant function coefficients					
System capability	0.07	−0.12	−0.02	0.69	0.55
Use of techniques	−0.18	−0.06	0.15	0.02	−0.51
Attention to internal facets	0.39	0.28	0.29	−0.46	0.15
Attention to external facets	0.22	0.66	0.69	0.24	0.09
Functional coverage	0.07	0.24	0.10	0.56	0.45
Resources provided for planning	0.76	0.67	0.42	0.73	0.66
Resistance to planning	0.06	0.50	0.23	1.02	0.39

[a] The failure of Box's M test for equality of group dispersion matrices suggests that quadratic classification procedures might be more appropriate in three of the analyses. The use of quadratic classification did not, however, lead to a noticeable improvement in classification accuracies.

satisfaction variable, as expected. The generally low and nonsignificant correlations between the objective fulfillment and relative performance variables does confirm our initial treatment of the two as distinct criteria of effectiveness. Thus, the uniformly good discrimination achieved across all the runs attests to the value of our seven dimensions of planning systems in classifying planning systems along either type of effectiveness criterion.

Equality of group sizes is an important consideration in interpreting bias in discriminant analysis (Morrison, 1969). The grouping procedure used in this study yielded nearly equal groups, with the smaller groups accounting for no less than 45 percent for all but the runs based on alternative evaluation, long-term performance, growth in sales, and growth in earnings. For these runs, the accuracies of chance models used as benchmarks are accordingly higher. In other words, we compared most of our discriminant functions' accuracies in classification against relatively sound test values in judging their worth. As can be noted in Table 6, the most unequal group sizes resulted for evaluating alternatives. Even in this case, the discriminant function's overall classification accuracy of 66.7 percent is somewhat better than the 63.9 percent expected accuracy of the corresponding chance model. For growth in earnings, the discriminant function's overall classification accuracy was about 9 percentage points better than that of the corresponding chance model. In all other cases the overall classification accuracy of the discriminant functions was at least 15 percentage points higher than that of the corresponding chance models. Thus, the results of the discriminant analysis runs are informative.

Traditionally, scholars have believed that the absolute magnitudes of the standardized discriminant function coefficients of discriminating variables reflect their relative importance. These magnitudes have been considered to be analogous to the beta weights in a regression model (Klecka, 1975). However, some methodologists question this approach. Although many alternative ranking schemes exist, this study used the procedure suggested by Mosteller and Wallace (1963) for ranking the discriminating dimensions.[4] Table 8 summarizes the relative importance rankings for the 13 analyses.

Satisfaction with Planning Systems

For the grouping based on satisfaction with planning systems, the three top-ranked discriminators were resistance to planning, system capability, and resources provided for planning. Attention to internal facets ranked as the least important discriminating dimension. These findings suggest that satisfaction with planning is largely a function of its organizational context.

On a univariate basis, the low-satisfaction group differed from the high-satisfaction group along all but one dimension: attention to external facets.

---

[4] In this procedure, the variables are ranked on the basis of the absolute value of the product $bj\ (\bar{X}_{j1} - \bar{X}_{j2})$, where bj is the unstandardized discriminant function coefficient of the variable j, $\bar{X}_{j1}$ is the mean of this variable for group 1, and $\bar{X}_{j2}$ is the mean for group 2. For discussion of the superiority of this procedure over the conventional use of the standardized discriminant coefficients for ranking, see Mosteller and Wallace (1963) or Joy and Tollefson (1975).

TABLE 8
Relative Importance Rankings of the Dimensions of Planning Systems in 13 Discriminant Analyses

	Dimensions						
Effectiveness Measures	System Capability	Use of Techniques	Attention to Internal Facets	Attention to External Facets	Functional Coverage	Resources Provided for Planning	Resistance to Planning
Objective fulfillment							
Predicting future trends	2	1	5	7	3	6	4
Evaluating alternatives	1	6	4	5	7	3	2
Avoiding problem areas	1	3	2	7	6	5	4
Enhancing management development	1	7	5	6	2	4	3
Improving short-term performance	7	1	6	4	5	2	3
Improving long-term performance	6	5	7	3	4	1	2
Objective composite	1	4	5	6	3	7	2
Relative performance							
Growth in sales	5	4	2	3	7	1	6
Growth in earnings	6	7	5	1	3	2	4
Changes in market share	7	5	3	1	6	2	4
Return on investment	3	7	6	5	2	1	4
Performance composite	2	5	6	7	3	1	4
Satisfaction							
Satisfaction with planning system	2	6	7	4	5	3	1

However, in the discriminant analysis, attention to external facets emerged as the fourth most important discriminator.

If use of techniques, attention to internal facets, and functional coverage can be considered to represent aspects of a planning process that tap its degree of formality, then it appears that formality of planning and satisfaction are unrelated or only weakly related.

The implications of these findings for managers concerned with enhancing satisfaction with planning would appear to be to pay specific attention to the organizational context of planning, particularly to any sources of resistance to planning, and to the level of resources committed to planning activities. Enhancing the external orientation of a planning system also appears to contribute to satisfaction with planning. Presumably, organizational members perceive externally oriented planning systems as conducive to meaningful planning exercises and therefore such an orientation increases satisfaction with planning. Of course, these are conjectures subject to further empirical testing, since the results of discriminant analysis can only yield tentative clues as to causal linkages.

Fulfillment of Objectives

Turning to the analyses for the variables measuring fulfillment of objectives, system capability appears to be the key discriminator overall. It ranked first in four of the seven analyses of objective fulfillment. However, it ranked last in the analysis for short-term performance and sixth for long-term performance. Thus, system capability appears to be more related to the process-based as opposed to the outcome-based objectives of planning. Further support for this conjecture is provided in the fact that system capability ranked near the bottom in three of the five analyses based on the variables measuring relative performance. But it ranked second when the grouping for discriminant analysis was based on the performance composite (Table 8).

Although resistance to planning never ranked as the top discriminator in the analyses of objective fulfillment, it never ranked lower than fourth, and in five analyses it ranked either second or third. Thus resistance to planning appears to be a key discriminator for fulfillment of objectives.

Use of techniques, functional coverage, and resources provided for planning ranked among the top three discriminators in three analyses each. Each ranked at the bottom in one. Although use of techniques was top-ranked in the analyses predicting future trends and short-term performance, it ranked last in that for enhancing management development. Thus, the contributions of techniques seem limited to the relatively more tangible objectives. Functional coverage ranked second in the analysis for enhancing management development and third in those predicting future trends and the objective composite. Why functional integration should improve prediction of future trends is not clear, but it is intuitively tenable that such integration can lead to management development. Functional emphasis and integration can result in the identification of key success factors and in developing a general management perspective. Such integration can also be expected to reduce conflict across functional boundaries. Resources provided for planning was found to be an important discriminator for the performance objectives and for stimulating a wide variety of alternatives and their evaluation.

Attention to internal facets as well as attention to external facets appear to be the least important discriminators as far as the variables measuring fulfillment of objectives are concerned. However, attention to external facets ranked relatively high with reference to the performance objectives, and attention to internal facets achieved its highest ranking in the analyses for measures of avoiding problem areas and evaluating alternatives.

Relative Performance

Resources provided for planning emerged as the key discriminator in the analyses of relative performance. It ranked consistently as the first or second factor. In contrast, use of techniques had consistently low rankings. It ranked last in two analyses, namely those for growth in earnings and return on investment, and in no analysis did it rank higher than fourth. This finding lends support to recent criticisms of the technique orientation in contemporary management (Peters & Waterman, 1982), and particularly to the attacks against popular planning techniques such as PIMS and BCG (Kiechel, 1982; Wensley, 1982). Techniques have a role to play in strategic planning, but the results of this study suggest that effective planning requires more than merely using sophisticated analytical techniques.

The consistent and relatively low ranking of resistance to planning in the analyses for relative competitive performance is also an interesting finding. Early writings on planning placed considerable emphasis on the resistance issue. It is tempting to speculate that such emphasis, although appropriate in the early days of planning, is perhaps less critical today. Today planning has increasingly been recognized and accepted as a line manager's job (*Business Week*, 1984). The issue of introducing it to organizations and securing the commitment of managers may have simply become less crucial. This does not, however, miminize the need for managing organizational resistance to planning.

Attention to external facets ranked at the top in the analyses for growth in earnings and changes in market share, and was third in the analysis for growth in sales. Thus, an external orientation appears to be important for effectiveness in relative competitive performance. However, attention to external facets was not a key discriminator for return on investment and the performance composite.

Both attention to internal facets and functional coverage showed considerable instability in their rankings from analysis to analysis. The former's high ranking in the analysis for growth in sales is not easy to interpret; the high ranking of functional coverage in the analysis for return on investment is explainable as possibly due to the efficiency achieved by good functional coordination.

A Synthesis

Considering all analyses together, system capability and resources provided for planning each ranked at the top in four analyses. System capability ranked as one of the three top discriminators in eight analyses and resources in nine. System capability ranked at the bottom in only two analyses and resources in only one. Thus, overall there is little doubt that system capability and resources provided for planning are key dimensions when the effectiveness of planning systems is construed in the broadest sense.

Although it ranked at the top in only one analysis, resistance to planning was never ranked at the bottom and was among the top three discriminators in six of the analyses. However, its discriminating influence was limited to the analyses of objective fulfillment and satisfaction. In contrast, functional coverage, which also ranked among the top three discriminators in six analyses, was found to be an important discriminator in three of the seven analyses of fulfillment of objectives and three of the five analyses of relative performance. We are therefore inclined to consider functional coverage to be a more important discriminator than resistance to planning overall.

IMPLICATIONS

The importance of the context of planning is underscored by the fact that each of the three criteria of effectiveness employed in this study appears to be influenced significantly by at least one of the two contextual dimensions. Both resources provided for planning and resistance to planning were key discriminators for satisfaction as well as objective fulfillment, whereas resources ranked consistently at the top for relative competitive performance. Future studies should include more contextual dimensions than the two used in this study. Possible candidates include environmental volatility, leadership styles, and maturity of planning systems.

Among the design elements, system capability proved to be a powerful discriminator for satisfaction and objective fulfillment, but proved less important for relative performance—except for the performance composite. In contrast, attention to external facets was an important discriminator for relative performance, but not for objective fulfillment. Functional coverage was an important discriminator for both objective fulfillment and relative performance. These results suggest that the importance of design elements varies considerably depending on the criterion of effectiveness used. The implication of this conclusion for systems design is that efforts to create supersystems capable of achieving numerous purposes at the same time are unlikely to be productive and that systems design should be tailored to specific purposes of planning. Trade-offs are inevitable if a planning system is to meet multiple aims (King, 1983).

The implication of the overall importance of resources and functional coverage is straightforward. Effective planning requires an infusion of adequate resources to planning efforts. This in turn requires high levels of involvement on the parts of not only staff managers but also operating managers and chief executive officers. As for functional coverage, a high score along this dimension would result from uniformly high responses on all the component items in this scale. This implies that functional integration and a general management perspective are crucial for effective planning.

The finding that system capability emerged overall as the most important discriminator raises important questions about the roles of creativity and control as key considerations in the design of planning systems. Some component items of this dimension tapped an orientation toward creativity and others tapped an orientation toward control. Although it is conventional to think of creativity and control as opposed orientations, our data suggested that the two tend to go together, as Peters and Waterman implied in their discussion of "simultaneous loose-tight properties" (1982: 318–325) in the context of organizational excellence. This was the reason we included aspects of creativity and control within the same dimension. For the present, we suggest that effective planning is achieved by a high degree of attention to both creativity and control, as opposed to trading the one off against the other. We are currently subjecting this hypothesis to further empirical analysis by examining interactive models of the creativity and control subdimensions of the system capability construct.

CONCLUSIONS

Most prior research on the effectiveness of planning systems has pursued the implicit question, "Does planning pay?" In contrast, this study was based on the premise that planning is a multidimensional and indistinguishable feature of overall management processes, and hence we pursued the question, "What dimensions of planning are associated with effectiveness as approached from multiple perspectives?"

The results of this study suggest that determinants of the effectiveness of planning systems tend to vary depending on the specific criterion of effectiveness used. Key dimensions overall seem to be system capability, resources provided for planning, and functional coverage.

Every effort was made to use an appropriate methodology and to execute it as rigorously as possible. However, a key limitation of this study is our reliance on data from a single respondent in each firm. This may have introduced response or function bias to the data, and it may raise questions about the generalizability of our findings, since organizational constructs like these are probably best measured with data from multiple respondents drawn from different functional and hierarchical levels. Also, we cannot be sure there was no bias resulting from nonresponse and incomplete data.

There are several possible extensions of this study. One line of inquiry would be to examine in detail the propositions suggested by the results summarized in Table 8. A second extension would be to examine the interrelationships among the dimensions of planning systems, instead of treating them as separate entities as we did. Third, future research could consider the seven systems dimensions jointly to develop rich empirically based taxonomies of planning systems, and their implications for effectiveness could be further explored. Lastly, future research could simultaneously analyze the relationships between the seven dimensions and the three criteria of effectiveness using approaches such as multivariate analysis of variance and canonical correlation analysis (Ramanujam & Venkatraman, 1985).

We believe that research on planning systems is at a critical crossroads. We hope this study will stimulate further efforts in new directions and enhance understanding of the role and value of planning systems in the formulation and implementation of strategies.

REFERENCES

Aguilar, F. J. 1965. ***Scanning the business environment.*** New York: Macmillan Book Publishing Co.

Amara, R. C. 1981. The futures field: Searching for definitions and boundaries. ***The Futurist,*** 15 (1): 25–29.

Andrews, K. R. 1971. ***The concept of corporate strategy.*** Homewood, Ill.: Dow Jones-Irwin.

Ansoff, H. I. 1965. ***Corporate strategy.*** New York: McGraw-Hill Book Co.

Ansoff, H. I. 1975. Managing strategic surprise by response to weak signals. ***California Management Review,*** 18 (2): 21–33.

Ansoff, H. I. 1984. ***Implanting strategic management.*** Englewood Cliffs, N.J.: Prentice-Hall.

Ansoff, H. I., Avner, J., Brandenburg, R. G., Portner, F. E., & Radosevich, R. 1970. Does planning pay? The effect of planning on the success of acquisitions in American firms. ***Long Range Planning,*** 3 (2): 2–7.

Anthony, R. N., & Dearden, J. 1976. ***Management control systems.*** Homewood, Ill.: Richard D. Irwin.

Armstrong, J. S. 1982. The value of formal planning for strategic decisions: Review of empirical research. ***Strategic Management Journal,*** 3: 197–211.

Bock, R. D. 1975. ***Multivariate statistical methods in behavioral research.*** New York: McGraw-Hill.

Business Week. 1984. The new breed of strategic planner. No. 2860 (September 17): 62.

Camillus, J. C. 1975. Evaluating the benefits of formal planning. ***Long Range Planning,*** 8 (3): 33–40.

Camillus, J. C., & Venkatraman, N. 1984. Dimensions of strategic choice. ***Planning Review,*** 12 (1): 26–31, 46.

Eisenbeis, R. A. 1977. Pitfalls in the application of discriminant analysis in business, finance, and economics. ***Journal of Finance,*** 32: 875–900.

Fahey, L., & King, W. R. 1977. Environmental scanning for corporate planning. ***Business Horizons,*** 20 (4): 61–71.

Frank, R. E., Massy, W. M., & Morrison, D. G. 1965. Bias in multiple discriminant analysis. ***Journal of Marketing Research,*** 2: 250–258.

Fredrickson, J. W. 1984. The comprehensiveness of strategic decision processes: Extensions, observations, future directions. ***Academy of Management Journal,*** 27: 445–466.

Fulmer, R. M., & Rue, L. W. 1974. The practice and profitability of long range planning. ***Managerial Planning,*** 22 (6): 1–7.

Gaedeke, R. M., & Tootelian, D. H. 1976. The Fortune "500" list: An endangered species for academic research. ***Journal of Business Research,*** 4: 283–288.

Gluck, F., Kaufman, S., & Walleck, A. S. 1980. Strategic management for competitive advantage. ***Harvard Business Review***, 58 (4): 154–161.

Gluck, F., Kaufman, S., & Walleck, A. S. 1982. The four phases of strategic management. ***Journal of Business Strategy***, 2 (3): 9–21.

Grant, J. H., & King, W. R. 1979. Strategy formulation: Analytical and normative models. In D. E. Schendel & C. W. Hofer (Eds.), ***Strategic management: A new view of business policy and planning***: 104–122. Boston: Little, Brown & Co.

Grant, J. H., & King, W. R. 1982. ***The logic of strategic planning***. Boston: Little, Brown & Co.

Grinyer, P. D., & Norburn, D. 1975. Planning for existing markets: Perceptions of executives and financial performance. ***Journal of the Royal Statistical Society***, 138 (Series A): 70–97.

Hax, A. C., & Majluf, N. S. 1984. ***Strategic management: An integrative perspective***. Englewood Cliffs, N. J.: Prentice-Hall.

Herold, D. M. 1972. Long range planning and organizational performance: A cross-validation study. ***Academy of Management Journal***, 15: 91–104.

Hitt, M. A., Ireland, R. A., & Palia, K. A. 1982. Industrial firms' grand strategy and functional importance: Moderating effects of technology and structure. ***Academy of Management Journal***, 25: 265–298.

Hitt, M. A., Ireland, R. A., & Stadter, G. A. 1982. Functional importance and company performance: Moderating effects of grand strategy and industry type. ***Strategic Management Journal***, 3: 315–330.

Hofer, C. W., & Schendel, D. E. 1978. ***Strategy formulation: Analytical concepts***. St. Paul, Minn.: West Publishing Company.

Joy, O. M., & Tollefson, J. O. 1975. On the financial applications of discriminant analysis. ***Journal of Financial and Quantitative Analysis***, 10: 723–739.

Karger, D. W., & Malik, Z. A. 1975. Long range planning and organizational performance. ***Long Range Planning***, 8 (6): 60–69.

Keegan, W. J. 1974. Multinational scanning: A study of information sources utilized by headquarters executives in multinational companies. ***Administrative Science Quarterly***, 19: 411–421.

Kefalas, A., & Schoderbek, P. P. 1973. Scanning the business environment: Some empirical results. ***Decision Sciences***, 4: 63–74.

Kiechel, W. 1982. Corporate strategists under fire. ***Fortune***, 106 (13): 34.

King, W. R. 1983. Evaluating strategic planning systems. ***Strategic Management Journal***, 4: 263–277.

King, W. R., & Cleland, D. I. 1978. ***Strategic planning and policy***. New York: Van Nostrand Reinhold Co.

Klecka, W. R. 1975. Discriminant analysis. In N. H. Nie, C. H. Hull, J. G. Jenkins, K. Steinbrenner, & D. H. Bent (Eds.), ***SPSS: Statistical package for the social sciences***: 434–467. New York: McGraw-Hill Book Co.

Kudla, R. J. 1980. The effects of strategic planning on common stock returns. ***Academy of Management Journal***, 23: 5–20.

Lindsay, W. M., & Rue, L. W. 1980. Impact of organization environment on the long range planning process: A contingency view. ***Academy of Management Journal***, 23: 385–404.

Lorange, P. 1979. Formal planning systems: Their role in strategy implementation. In D. E. Schendel & C. W. Hofer (Eds.), ***Strategic management: A new view of business policy and planning***: 226–241. Boston: Little, Brown & Co.

Lorange, P. 1980. ***Corporate planning: An executive viewpoint***. Englewood Cliffs, N.J.: Prentice-Hall.

Lorange, P., & Vancil, R. F. 1977. ***Strategic planning systems***. Englewood Cliffs, N. J.: Prentice-Hall.

Lucas, H. C. 1978. Empirical evidence for a descriptive model of implementation. ***MIS Quarterly***, 2 (2): 27–41.

Malik, Z. A., & Karger, D. W. 1975. Does long range planning improve company performance? ***Management Review***, 64 (9): 26–31.

Mintzberg, H. 1981. What is planning anyway? ***Strategic Management Journal***, 2: 319–324.

Morrison, D. M. 1969. On the interpretation of discriminant analysis. ***Journal of Marketing Research***, 6: 156–163.

Mosteller, F., & Wallace, D. F. 1963. Inference in an authorship problem. ***Journal of the American Statistical Association***, 58: 275–309.

Nunnally, J. C. 1978. ***Psychometric theory***. New York: McGraw-Hill Book Co.

Paul, R. N., Donavan, N. B., & Taylor, J. W. 1978. The reality gap in strategic planning. ***Harvard Business Review***, 56 (4): 124–130.

Peters, T. J., & Waterman, R. H. 1982. ***In search of excellence: Lessons from America's best run companies***. New York: Harper & Row Publishers.

Ramanujam, V., & Venkatraman, N. 1985. ***The assessment of strategic planning effectiveness: A canonical correlation approach***. Paper presented at the TIMS/ORSA Joint National Meeting, Boston.

Robinson, R. B., & Pearce, J. A. 1984. Research thrusts in small firm strategic planning. ***Academy of Management Review***, 9: 128–137.

Rue, L. W., & Fulmer, R. W. 1973. Is long range planning profitable? ***Academy of Management Proceedings***, 66–73.

Schultz, R. L., & Slevin, D. P. 1976. ***Implementation of OR/MS***. New York: Elsevier-North Holland.

Shank, J. K., Niblock, E. G., & Sandalls, W. T., Jr. 1973. Balance "creativity" and "practicality" in formal planning. ***Harvard Business Review***, 51 (1): 87–95.

Shrivastava, P., & Grant, J. H. 1985. Empirically derived models of strategic decision-making processes. ***Strategic Management Journal***, 6: 97–113.

Snow, C. C., & Hrebiniak, L. G. 1980. Strategy, distinctive competence, and organizational performance. ***Administrative Science Quarterly***, 25: 317–336.

Snyder, N. H. 1982. What is planning anyway? A rejoinder. ***Strategic Management Journal***, 3: 265–268.

Steiner, G. A. 1979. ***Strategic planning: What every manager must know***. New York: Free Press.

Steiner, G. A., & Schollhammer, H. 1975. Pitfalls of multinational strategic planning. ***Long Range Planning***, 8 (2): 2–12.

Stevenson, H. H. 1976. Defining corporate strengths and weaknesses. ***Sloan Management Review***, 17 (3): 51–68.

Thomas, P. S. 1980. Environmental scanning: The state of the art. ***Long Range Planning***, 13 (1): 20–28.

Thompson, J. D. 1967. ***Organizations in action***. New York: McGraw-Hill Book Co.

Thune, P. S., & House, R. 1970. Where long range planning pays off. ***Business Horizons***, 13 (4): 81–87.

Venkatraman, N., & Grant, J. H. 1986. Construct measurement in organizational strategy research: A critique and proposal. ***Academy of Management Review***, 11: 71–87.

Wensley, R. 1982. PIMS and BCG: New horizons or false dawn? ***Strategic Management Journal***, 3: 147–158.

Wood, D. R., & LaForge, R. W. 1979. The impact of comprehensive planning on financial performance. ***Academy of Management Journal***, 22: 516–526.

APPENDIX

Responses for all items were measured on 5-point scales. Items followed by (R) were reverse-coded to indicate resistance.

Dimensions of planning systems. System capability was measured on a scale ranging from "much improvement" to "much deterioration" with the following 12 items: ability to anticipate surprises and crises; flexibility to adapt to unanticipated changes; value as a mechanism for identifying new business opportunities; role in identifying key problem areas; value as a tool for managerial motivation; capacity to generate new ideas; ability to communicate top management's expectations down the line; value as a tool for management control; capacity to foster organizational learning; ability to communicate line management's concerns to top management; value as a mechanism for integrating diverse functions and operations; and value as a basis for enhancing innovation. In addition, the following item, scaled from "strongly disagree" to "strongly agree," was included: "Today's system emphasizes creativity among managers more than our previous system."

Use of techniques was measured on a scale ranging from "significant decrease in use" to "significant increase in use" with the following nine items: portfolio (e.g., BCG) approaches, PIMS models, financial models, zero-based budgeting, value-based planning, project management techniques (e.g., PERT/CPM), stakeholder analysis, scenarios/delphi techniques, and forecasting and trend analysis.

Attention to internal facets was measured on a scale ranging from "significantly less emphasis" to "significantly more emphasis" with the following three items: internal capabilities, past performance, and reasons for past failures.

Attention to external facets was measured on a scale ranging from "significantly less emphasis" to "significantly more emphasis" with the following five items: general economic and business conditions, regulatory issues, worldwide competitive trends, supplier trends, customers and end user preferences, and technological trends.

Functional coverage was measured on a scale ranging from "significantly less emphasis" to "significantly more emphasis" with the following seven items: marketing function, operations and manufacturing function, finance function, personnel function, purchasing and procurement function, R&D and technology, and computers and MIS.

Resources provided for planning was measured on a scale ranging from "significant decrease" to "significant increase" with the following four items: number of planners, time spent by the chief executive officer in strategic planning, involvement of staff managers in strategic planning, and resources provided for strategic planning.

Resistance to planning was measured on a scale ranging from "significant decrease" to "significant increase" with the following five items: overall emphasis on strategic planning (R), involvement of line managers in strategic planning (R), acceptance of the outputs of strategic planning exercise by top management (R), resistance to planning in general, and threats to the continuation of strategic planning.

Effectiveness of planning systems. Fulfillment of objectives over the past five years was measured on a scale ranging from "entirely unfulfilled" to "entirely fulfilled" with the following six items: predicting future trends, evaluating alternatives based on more relevant information, avoiding problem areas, enhancing management development, improvement in short-term performance, and improvement in long-term performance.

Performance relative to competition over the past five years was measured on a scale ranging from "much worse" to "much better" with the following four items: growth in sales, growth in earnings, changes in market share, and return on investment.

Vasudevan Ramanujam is an assistant professor in management policy at the Weatherhead School of Management, Case Western Reserve University. He received his Ph.D. degree in strategic planning and policy from the University of Pittsburgh. His current areas of research interest include strategic planning systems, corporate turnaround, and technological innovation and change.

N. Venkatraman is an assistant professor at the Alfred P. Sloan School of Management, Massachusetts Institute of Technology. He earned his Ph.D. degree in strategic planning and policy from the University of Pittsburgh. His current research interests include strategic planning systems, construct measurement, and the use of methodologies for modeling structural equations in strategy research.

John C. Camillus is Associate Dean and an associate professor at the Graduate School of Business, University of Pittsburgh. He received his D. B. A. degree from Harvard University. His current research focuses on strategic planning and management control systems.

© *Academy of Management Journal*
1986, Vol. 29, No. 2, 373–384.

LAYOFFS, EQUITY THEORY, AND WORK PERFORMANCE: FURTHER EVIDENCE OF THE IMPACT OF SURVIVOR GUILT

JOEL BROCKNER
Columbia University
JEFF GREENBERG
University of Arizona
AUDREY BROCKNER
Hunter College
JENNY BORTZ
JEANETTE DAVY
CAROLYN CARTER
University of Arizona

This study explored the effects of layoffs on survivors. We assessed subjects' work performance as a function of whether a co-worker had been laid off and the circumstances of that layoff. Consistent with equity theory, subjects worked harder when they believed that a co-worker's dismissal was based on a random process rather than on the relative merits of their and their co-worker's prior performance. Data drawn from questionnaires lent further support to equity theory as an explanation of these results.

Many organizations resort to layoffs during periods of financial strain in obvious attempts to reduce personnel costs. Less obvious though perhaps no less important are the likely consequences that layoffs have for the work behavior and attitudes of survivors, those employees who are not laid off. To date relatively little research has explored the effects of layoffs on survivors. The present study was designed to address this deficiency.

Several psychological perspectives offer some insight into the effects of layoffs on survivors. For example, layoffs may elicit anxiety produced by job insecurity that may in turn influence the performance of survivors. Indeed, the relationship between anxiety and performance has long been a topic of interest to experimental psychologists (Spence, Farber, & McFann, 1956).

More germane to the present study, the voluminous literature on distributive justice (Deutsch, 1975; Folger & Greenberg, 1985; Leventhal, 1980; Tyler & Caine, 1984) may also elucidate the effects of layoffs on survivors. Distributive justice concerns the principles decision makers rely on in

---

The authors thank Max Bazerman, David Lewin, Mike O'Malley, Lance Sandelands, Mike Tushman, and two anonymous reviewers for their helpful comments on an earlier version of the manuscript.

allocating scarce resources. According to Deutsch and others, the relative salience of these principles depends on the goals that a group or organization is trying to achieve. In business organizations the primary goal is economic productivity; thus equity—rather than equality or need—is most likely to be the salient norm. Accordingly, the vast body of research on equity theory (Adams, 1965) may apply to the current analysis.

Much of the previous research on equity theory in industrial and organizational psychology has evaluated the effects of positive inequity on workers' reactions (Mowday, 1979). Positive inequity refers to workers' perceiving that the ratio between outcome and input is higher for them than it is for relevant others. Two basic notions in equity theory are that positive inequity (1) arouses guilt, and (2) motivates individuals to redress this guilt through behavioral or psychological means. For example, Adams (1965) and others have shown that positive inequity may cause employees to work harder, presumably in order to redress the perception that they are taking advantage of their employers.

It is our contention that layoffs may cause survivors to experience positive inequity, especially if they believe that they could have just as easily been dismissed as their co-workers. Thus, contrary to the procedure employed in most studies of positive inequity, which create positive inequity by increasing workers' outcomes in relation to those of comparison others, surviving a layoff may produce positive inequity by decreasing outcomes for comparison others.

In fact, a recent study by Brockner, Davy, and Carter (1985) tested the hypothesis that the dismissal of a co-worker could cause survivors to experience guilt produced by positive inequity that in turn would lead to improvement in subsequent work performance. In that laboratory experiment, each subject performed a clerical task and was led to believe that another person—who was actually an experimental accomplice—was simultaneously performing a similar task. Subjects' motivation for taking part in the study was to receive extra course credit. Midway through the proceedings, half of the subjects were told that an unforeseen problem made it impossible for both them and the other person to complete the experiment. Moreover, they were told that the person who did not complete the experiment would receive no credit for having taken part in the study. A seemingly random procedure always forced the accomplice to leave while the subjects survived. These subjects, whom we designated as in the layoff condition, then performed the second half of the clerical task.

The other half of the subjects, designated as in the control condition, were merely interrupted midway through the proceedings for a break. The primary dependent variables were (1) subjects' perceptions of how fairly both they and the other person had been treated, (2) their guilt, and (3) their work performance, as indicated by the quantity of their productions and quality of their efforts on the first and second halves of the clerical task.

The results were consistent with equity theory. Compared to those in the control condition, subjects in the layoff condition (1) reported that they had been treated much more fairly than the other person, (2) reported feeling more guilty, and (3) increased the quantity—but not the quality—of their performance to a greater extent on the second half of the task.

In this previous study (Brockner et al., 1985), decisions about whom to lay off were results of a seemingly random process. For practical as well as theoretical reasons, however, it seems important to study the effects of different rules for making layoff decisions on the work performance of survivors. At the practical level, managers faced with the difficult task of implementing layoffs will clearly be concerned with whether different decision rules elicit differential responses among survivors. If a particular decision rule causes survivors to feel and perform more positively, or at least less negatively, managers may wish to invoke that rule when layoffs are inevitable. At the theoretical level, equity theory provides reason to believe that decision rules significantly affect the subsequent performance of survivors.

The present study was thus designed to extend the earlier study. We randomly assigned subjects to one of three experimental conditions. As in the previous study, we studied random layoff and control conditions. In addition, we assigned a third group of subjects to a merit layoff condition in which they were told that the decision about who was to stay was based on the relative merit of their performance on the first half of the task. We led subjects in the merit layoff condition to believe that they had outperformed an accomplice.

Equity theory predicts that the merit layoff condition should have less effect on workers' subsequent performance than will the random layoff condition. More specifically, subjects in the merit layoff condition may restore equity *psychologically* by perceiving that their better performance or greater input justified their status as survivors. Consequently, these survivors have less need to redress the inequity behaviorally through such means as heightened work performance. In short, although we expected the quantity of survivors' output to be greater in the random than in the control condition, we expected that there would be little or no difference in quantity of output between those in the merit and control conditions. Instead, subjects in the merit condition would redress inequity by focusing on the fact that they had performed better than their co-worker on the previous task.

The present study also addressed another important theoretical issue. At the outset we implied that guilt produced by positive inequity is one, but not the only, reaction that survivors may have to co-workers' layoffs. In actual organizations, survivors may have many other feelings. To name two possibilities, they may be worried or anxious that they too will be laid off at some future date, or angry, especially if they perceive the layoffs as illegitimate. Because the present study is derived from equity theory, it is important to isolate the effects of guilt produced by positive inequity from these other reactions of survivors. Thus, we induced the layoff conditions so as to arouse guilt produced by positive inequity but not other affective states like anxiety or anger. Of course, whether subjects in the layoff conditions actually felt anxious or angry in addition to feeling guilty is an empirical question. Accordingly, we assessed their affective states after the layoff manipulation.

METHODS

Subjects

A total of 145 undergraduates at a large state university completed the experiment. We drew subjects from introductory psychology or management classes, offering them extra course credit for taking part.

Procedures

Upon arrival at the laboratory for what we called a test validation experiment, a subject was led into a large research room and seated at a desk. The experimenter told the individual that another subject had to arrive before the experiment would begin. Approximately 30 seconds later a woman confederate, posing as another subject, arrived and was seated at a desk diagonally opposite to the subject. Both subject and confederate were informed that they would (1) work on a proofreading task, (2) have a short break, and (3) work on a second proofreading task. The alleged purpose of the proofreading was to develop some standardized tests of cognitive performance. To minimize competition, we told both individuals that they would not be proofreading the same material. Furthermore, they were told that "to minimize any possible distractions we are going to have you do the proofreading in separate, private rooms."

At this point we ushered both confederate and subject into separate research cubicles to perform a proofreading task that has been used in previous research (Reis & Burns, 1982). They received two six-page booklets containing a passage from Jane Jacobs's *Death and Life of Great American Cities*. One copy was designated uncorrected pages and had numerous spelling, grammatical, and punctuation errors throughout. The other, flawless copy was designated master pages. Subjects were instructed to proofread by comparing the two and to underline but not correct all of the errors they observed on the uncorrected pages. They had eight minutes to read the passage. The experimenter emphasized the importance of both speed and accuracy to subjects; these provided indicators of quantity and quality of performance.

The experimenter then left subjects alone in their research cubicles for eight minutes to perform the proofreading. After eight minutes had elapsed the experimenter reentered subjects' cubicles, had them underline the last line they had proofread, and asked them to follow her back to the large research room. The confederate, also fetched by the experimenter, came to the large research room several seconds later. During this rest period we introduced the layoff manipulation. Previously we had randomly assigned subjects to one of the three conditions. The experimenter told subjects in the random and merit layoff conditions that because of a room scheduling problem only one person could finish the experiment; moreover, the person who left would receive no credit for having taken part in the study. Then the experimenter stated an apparently impromptu basis on which the layoff decision was going to be made.

The experimenter told subjects in the random layoff condition that both individuals would have to draw lots to determine which one of them would get to stay. We rigged the lottery so that subjects always were chosen to stay. For subjects in the merit layoff condition the experimenter said that the decision would be based on the quantity and quality of each individual's performance on the first proofreading task. The experimenter then excused herself for a few moments, supposedly to grade the first proofreading tasks. Subjects were then always informed that they had performed better than the other person and therefore could stay while the confederate was forced to leave.

To buttress the perceived unfairness of her having been forced to leave, the confederate said as she was exiting that it was really unfair to make her leave and deprive her of the extra credit. The experimenter verbally agreed that it was a tough break, but said nothing else.

For the control condition the experimenter merely asked subjects not to speak with one another during the rest period. The experimenter then made busy work for herself, for about one minute, a period comparable to the time needed to complete the layoff for the random and merit layoff conditions.

Subjects in all conditions were led back to their private research cubicles, in which they worked on a different six-page proofreading task taken from the same book used for the first task. For the control condition the confederate was led back to her cubicle first, in full view of the subject. The instructions and length of time allotted for the second task were the same as for the first. After eight minutes the experimenter reentered the subjects' cubicles and asked them to underline the last line that they had proofread.

Dependent variables. The behavioral dependent measures were measures of the quantity and quality of proofreading performance on each task. We assessed the first by counting the numbers of lines that subjects had proofread during the eight-minute reading periods and determined the second by computing the percentages of errors that subjects had correctly identified in the lines they had finished reading.

After completing the second proofreading task all subjects responded to a final questionnaire, which included a manipulation check as well as a number of measures of their affective states while working on the second proofreading task. The two manipulation check questions required subjects to rate (1) how fairly the experimenter treated them (self-rating), and (2) how fairly the experimenter treated the other subject (other rating). Responses could range from 1 = not at all to 7 = more than fairly. Subjects also indicated the extent to which each of the following described how they felt while proofreading the second task: frustrated, worried, lucky, sad, confident, guilty, anxious, glad, competitive, angry, helpless, distracted, and sorry for the other person. Responses to each of these measures could range from 1 = not at all to 7 = a great deal. In addition, subjects responded to an item designed to tap their perceived inputs on the first proofreading task: "How well did you perform on the first proofreading task?"; responses could range from 1 = very poorly to 7 = very well.

After they had completed this questionnaire the experimenter fully debriefed the subjects, paid and thanked them, and implored them not to speak with anyone about the research.

RESULTS

Manipulation Check

Equity theory posits that perceived equity is the result of relative comparison—individuals' comparing themselves to others. Thus, we employed an analysis of variance with two factors, condition × self-other, with repeated measures on the second factor. The relevant effect is the interaction, which was highly significant ($F_{2,142} = 14.15, p < .001$). Subjects in the random layoff condition reported that they had been treated much more fairly than the other subject ($\bar{x} = 5.22$ and 3.44), and subjects in the merit layoff condition did likewise ($\bar{x} = 5.34$ and 3.63). But those in the control condition did not report that they had been treated more fairly ($\bar{x} = 6.08$ and 5.79). The layoff induction thus appears to have been successful. Moreover, there was no evidence that subjects saw layoffs based on merit as less unfair than the random layoffs. What differed was how the subjects in the two layoff conditions responded to the perceived positive inequity.

Behavioral Measures

Quantity of performance. Subjects' performance on the proofreading task prior to the layoff manipulation did not differ as a function of experimental condition ($F < 1$). We then performed a two-factor (condition × task) analysis of variance with repeated measures on the second factor. The main effect for task was highly significant ($F_{1,142} = 59.27, p < .001$). Across conditions, subjects proofread many more lines in the second than the first task. This effect, which was not of primary importance, may have been due to a practice or warm-up effect.

Of greater importance was the significance of the interaction between condition and task ($F_{2,142} = 3.90, p < .025$). As predicted and shown by Table 1, the greatest rise in productivity between the first and second tasks occurred in the random layoff condition ($\bar{x} = 15.55$ lines). By contrast, subjects in the control condition proofread 6.51 more lines on the second than the first task and those in the merit condition proofread 9.41 more lines on the second task. The results of a Newman-Keuls multiple comparison test revealed that the increase in the random condition was significantly greater than that exhibited in the control condition ($p < .02$), and in the merit condition ($p < .06$). As predicted, there was no significant difference between the increases exhibited in the control and merit conditions ($p > .25$). The significant difference between the random and control conditions thus replicates a previous research finding (Brockner et al., 1985) whereas the comparisons between

TABLE 1
Comparison of Productivity on Tasks for Three Experimental Conditions

	Mean Number of Lines Proofread[a]		
Experimental Conditions	**First Task**	**Second Task**	**Increase in Lines Proofread[b]**
Control	85.31	91.82	6.51
(N = 39)	(25.06)	(27.21)	
Random layoff	87.67	103.22	15.55
(N = 36)	(27.34)	(29.95)	
Merit layoff	84.07	93.48	9.41
(N = 70)	(25.72)	(23.48)	

[a] Maximum score = 164 lines; standard deviations in parentheses.

[b] The layoff manipulation took place between the first and second tasks.

the merit and random layoff conditions and the merit and control conditions extend the earlier research.[1]

On the assumption that work performance is a function of motivation and ability, it may be useful to speculate about what mediated the significant results on the quantity measure. It is likely that the main effect for task reflected an increase in proofreading ability due to a practice effect from the first to the second task. The varied increase in work quantity across the three conditions, however, more likely reflected differences in motivation than in ability.

Quality of performance. There were no differences in performance quality on the initial task as a function of the upcoming layoff manipulation ($F < 1$). We performed an analysis of variance for condition × task with repeated measures on the second factor. As in the analysis of the data on quantity, two significant effects emerged. First, the main effect for task was highly significant; subjects correctly identified fewer errors in the material in the second task ($F_{1,142} = 24.50, p < .001$). Second, and of greater importance, was the significant interaction between the factors ($F_{2,142} = 4.89, p < .01$). The greatest decline in quality between the first and the second tasks occurred for subjects in the random layoff condition ($\bar{x} = 63.50$ and 53.72%); the least decline occurred for those in the merit layoff condition ($\bar{x} = 61.03$ and 59.15%); and a moderate decline occurred in the control condition ($\bar{x} = 61.49$ and

---

[1] Brockner and colleagues (1985) observed an increase in productivity under the random layoff condition particularly among subjects who were low in chronic self-esteem. In order to evaluate replicability, we had subjects complete the Janis-Field self-esteem scale (Eagly, 1967) at the outset of the present study. An ANOVA for self-esteem × condition × task yielded no significant effects for self-esteem, but the effect of self-esteem in the control and random layoff conditions was quite similar to that observed in the previous study. A simple analysis of effects revealed that people with low self-esteem increased the quantity of their performance more in the random than in the control condition ($F_{1,69} = 5.31, p < .03$). In marked contrast, in the random and control conditions people with medium and high self-esteem exhibited equal increases in productivity (Fs < 1).

56.23%). The Newman-Keuls multiple comparison test revealed that the change in quality of performance differed between the random and the merit layoff condition ($p < .01$), but that changes in the two layoff conditions did not differ significantly from the change exhibited in the control condition.

Correlational analyses revealed that quantity and quality of performance were inversely related on both proofreading tasks, though somewhat more so on the second than the first ($r_{143} = -.32$ and $-.56$). In spite of these negative correlations, we should note that the data on change in quality between conditions were not simply a mirror image of the quantity results. Although subjects in the random condition exhibited both the sharpest decrease in quality and the sharpest increase in quantity, the condition in which quality of performance showed the least decline was the merit condition and the one in which the increase in quantity was the least was the control condition.

Affective States

Data from the manipulation check suggested that subjects in the layoff condition did perceive positive inequity in how they had been treated compared to their dismissed co-worker. Was this positive inequity associated with heightened guilt or remorse as equity theory would predict? We found that it was. Subjects in the merit and random layoff conditions reported feeling more guilty ($\bar{x} = 2.84$ and 2.36) while performing the second proofreading task than did subjects in the control condition ($\bar{x} = 1.72$; $F_{2,142} = 4.65$, $p < .01$). Furthermore, subjects reported feeling much more sorry for the other person in the merit and random layoff conditions ($\bar{x} = 4.81$ and 5.03) than in the control condition ($\bar{x} = 1.74$; $F_{2,142} = 35.73$, $p < .001$).

Although this evidence suggests that surviving a layoff evoked feelings of guilt or remorse, it does not rule out the possibility that the layoffs aroused other feelings unrelated to equity theory. For example, subjects in either of the layoff conditions could have felt worried or anxious about the possibility of their also being asked to leave, or angry. It is important, however, that subjects in the random and merit conditions did not report feeling more worried ($F < 1$), anxious ($F < 1$), or angry ($F < 1$) than did subjects in the control condition. In fact, the layoff manipulation had no significant effect on subjects' feeling frustrated, lucky, confident, competitive, sad, helpless, or distracted, according to self-reports. The only other significant effect emerged for feeling glad ($F_{2,142} = 5.22$, $p < .01$). Not surprisingly, subjects in the merit and random layoff conditions reported feeling more glad ($\bar{x} = 4.51$ and 4.06) than did those in the control condition ($\bar{x} = 3.49$).

Perceived Input

The results for the manipulation check and measures of affective states revealed that subjects in the random and merit conditions reported feeling relatively equal levels of positive inequity, guilt, and remorse. Yet only those in the random condition significantly increased productivity. We predicted that those in the merit condition would be able to redress the positive inequity through the notion of their greater input. They would justify their survivor status by focusing on the fact that their initial performance or the task was better. In fact, subjects in the merit condition did evaluate their initial task performance more favorably ($\bar{x} = 5.09$) than did those in the random ($\bar{x} = 4.31$) and control conditions ($\bar{x} = 4.49$; $F_{2,142} = 4.42$, $p < .01$). Of course, such results do not prove that survivors in the merit condition redressed the inequity psychologically, but they are at least consistent with such an assertion.

DISCUSSION

In summary, data revealed that subjects saw the random and merit layoffs as more unfair to the other subject and that these layoffs aroused guilt and remorse. Perhaps most interesting, subjects appeared to redress the positive inequity in very different ways in the two layoff conditions. Random layoffs caused subjects to increase the quantity of their output; merit layoffs led subjects to perceive that they had performed more favorably on the first task than did their counterparts in the control or random conditions. This suggests that the merit survivors restored equity through psychological rather than behavioral means. Finally, the quality of survivors' performance was significantly greater in the merit than the random condition.

Explanations

The present study was generally derived from distributive justice theory (Deutsch, 1975) and grounded more specifically in equity theory (Adams, 1965). Taken together, the three findings of the present study provide strong support for equity theory.

First, data from the manipulation check revealed that subjects perceived their co-worker's dismissal to be unfair; in other words, subjects reported experiencing positive inequity as a result of a comparison with the worker who was laid off.

Second, the results of the measures of affective states demonstrated that subjects in the merit and random layoff conditions reported feeling guilty and remorseful, both believed to be emotional consequences of positive inequity.

Third, equity theory posits that inequity can be redressed through both behavioral and psychological means. Once again, our evidence supports the theory. Survivors in the random condition exhibited a greater increase in subsequent performance input than did those in the merit and control conditions. Survivors in the merit condition perceived their prior performance input as greater than did survivors in the other two conditions.

Alternative interpretations. Not only do the findings reported in the preceding paragraphs support equity theory, but other findings—primarily those drawn from the measures of affective states—are inconsistent with alternative interpretations. Thus, surviving layoff could have increased anxiety among subjects; or in a more positive vein, it could have increased confidence. Either could have affected their subsequent work performance.

For example, survivors in the random condition may have increased the quantity of their output to the extent that they experienced worry or anxiety produced by job insecurity. However, there was absolutely no evidence that subjects in either the random or merit conditions felt worry, anxiety, confidence, or any of a variety of emotions more acutely than did subjects in the control condition.

The present findings extend the range of organizational phenomena to which equity theory is relevant. One frequent criticism of previous equity studies is that researchers have relied too much on overpayment to induce positive inequity. In actual organizations a variety of factors, not all having to do with money, can stimulate inequity. This study is one of the few attempts to manipulate inequity through nonfinancial means (cf. Greenberg & Ornstein, 1983).

Future Research

To reiterate, only a few earlier studies (e. g., Greenhalgh, 1983) have empirically investigated the impact of layoffs on survivors. We hope that the results reported here will stimulate further work on this topic of practical and theoretical significance. To increase the external validity of such future research efforts, investigators should study the effects of layoffs in actual organizations. It is likely that actual layoffs evoke many affective states among survivors—anxiety, guilt, anger, and relief, to name a few—that may influence their work behaviors and attitudes in complex ways. Investigators need to give careful thought to the factors that moderate the effects of layoffs on survivors. The results reported here suggest that the rule by which layoff decisions are made is one relevant factor; but there are countless others. For example, layoffs are bound to have more emotional impact when (1) prior work interdependence between the laid-off workers and survivors is high, (2) the survivors and laid-off workers are similar professionally or personally, and (3) survivors cannot easily find comparable work elsewhere. In sum, we suspect that many types of factors moderate the effects of layoffs on survivors. These include characteristics of the work itself, such as interdependence; individual differences among survivors, such as differences in self-esteem; formal organizational arrangements like the decision rule upon which layoffs are based; informal factors like similarity of co-workers; and environmental conditions such as the job market for comparable work. Of course, research exploring the effects of these factors should be theoretically grounded, as was the present study, in which we derived the decision rule variable from equity theory.

In addition, future investigators should assess the effects of several factors that are often present in actual organizations but that did not exist in our experimental setting. For example, the longevity of the performance effect obtained in the present study remains to be empirically evaluated. In addition, the effects of anticipated layoffs are worthy of attention. Once again, we emphasize that it is important to study the impact of layoffs on survivors in more naturalistic settings. If anything, the laboratory procedure used in this study, in which relatively little was at stake and in which survivors and laid-off workers were unacquainted, may grossly underestimate the impact of layoffs on survivors in actual organizations.

REFERENCES

Adams, J. S. 1965. Inequity in social exchange. In L. Berkowitz (Ed.), ***Advances in experimental social psychology***, vol. 2: 267–299. New York: Academic Press.

Brockner, J., Davy, J., & Carter, C. 1985. Layoffs, self-esteem, and survivor guilt: Motivational, affective, and attitudinal consequences. ***Organizational Behavior and Human Decision Processes***, 36: 229–244.

Deutsch, M. 1975. Equity, equality, and need: What determines which value will be used as the basis of distributive justice? ***Journal of Social Issues***, 31: 137–149.

Eagly, A. H. 1967. Involvement as a determinant of responses to favorable and unfavorable information. ***Journal of Personality and Social Psychology***, monograph supplement, 7: 1–15.

Folger, R., & Greenberg, J. 1985. Procedural justice: An interpretive analysis of personnel systems. In K. Rowland & G. Ferris (Eds.), ***Research in personnel and human resource management***, vol. 3: 141–183. Greenwich, Conn.: JAI Press.

Greenberg, J., & Ornstein, S. 1983. High status job title as compensation for underpayment: A test of equity theory. ***Journal of Applied Psychology***, 68: 285–297.

Greenhalgh, L. 1983. Managing the job insecurity crisis. ***Human Resource Management***, 22: 431–444.

Leventhal, G. S. 1980. What should be done with equity theory? In K. J. Gergen, M. S. Greenberg, & R. H. Willis (Eds.), ***Social exchange: Advances in theory and research***: 27–55. New York: Plenum.

Mowday, R. T. 1979. Equity theory predictions of behavior in organizations. In R. M. Steers & L. W. Porter (Eds.), ***Motivation and work behavior***: 124–146. New York: McGraw-Hill Book Co.

Reis, H. T., & Burns, L. B. 1982. The salience of the self in response to inequity. ***Journal of Experimental Social Psychology***, 18: 464–475.

Spence, K. W., Farber, I. E., & McFann, H. H. 1956. The relation of anxiety (drive) level to performance in competitional and noncompetitional paired-associate learning. ***Journal of Experimental Psychology***, 52: 296–305.

Tyler, T. R., & Caine, A. 1981. The role of distributional and procedural fairness in the endorsement of formal leaders. ***Journal of Personality and Social Psychology***, 41: 642–655.

Joel Brockner earned his Ph. D. degree in personality and social psychology from Tufts University. He is currently an associate professor of management at the Graduate School of Business, Columbia University. His research interests include the consequences of organizational exits, conflict escalation, and self-esteem processes.

Jeff Greenberg received his Ph. D. degree in social psychology from the University of Kansas. He is currently an assistant professor of psychology at the University of Arizona. His research focuses on processes for maintaining self-esteem and on depression.

Audrey Brockner completed her M. S. W. degree at Simmons College. She is now a lecturer at the Hunter College School of Social Work. Her research interests are psychotherapy, clinical supervision, and self-psychology.

Jenny Bortz recently completed her undergraduate degree in business administration and psychology at the University of Arizona. Her research interests are personnel and human resource management.

Jeanette Davy is a doctoral candidate in the Department of Management at the University of Arizona. Her research interests focus on the effects of layoffs on survivors and labor unions.

Carolyn Carter is a doctoral candidate in educational psychology at the University of Arizona. Her research interests are psychotherapy and organizational behavior.

© *Academy of Management Journal*
1986, Vol. 29, No. 2, 385–404.

CLERICAL EMPLOYEES' PERCEPTIONS OF INTRAORGANIZATIONAL CAREER OPPORTUNITIES

JACQUELINE LANDAU
Tulane University
TOVE HELLAND HAMMER
Cornell University

This study investigated the determinants of perceived ease of movement within an organization among two groups of clerical employees. It also investigated the relationship between perceived ease of movement and both organizational commitment and intention to quit. Younger employees with shorter job tenure who felt that their personal characteristics matched organizational criteria for mobility and who received feedback from their supervisors perceived themselves as having the greatest opportunities for advancement. Employees who desired mobility but perceived no opportunities were less committed to their organizations than employees who perceived they had opportunities for advancement.

In one model of organizational careers (Milkovich, Anderson, & Greenhalgh, 1976), "the career construct is represented in two forms, actual and perceptual" (1976: 20). Actual careers have the "dual properties of (1) a patterned sequence of positions usually related in work content and (2) an ordered movement of individuals among these positions" (1976: 20). Perceived opportunity is an employee's interpretation of actual careers. Although there have been numerous studies on actual careers and mobility, perceived opportunities has been a neglected variable in empirical research. This concept has, however, been addressed from a theoretical perspective. Perceived ease of movement, which refers to employees' assessments of their opportunities to find better jobs, is a central factor in March and Simon's (1958) model of behavior at work. According to their theory, dissatisfied employees will explore alternative situations. If they see that alternatives exist—in other words, perceived ease of movement is high—these employees may leave their present situations. If perceived ease of movement is low because alternatives are not available, employees may adjust their aspirations or contributions to their organizations downward.

March and Simon (1958) were primarily interested in determinants of perceived ease of movement outside an organization. The focus of this study

---

The authors wish to thank Ed Kemery for advice on statistical analysis. An earlier version of this article was presented at the annual meeting of the Academy of Management in Dallas, 1983.

is perceived ease of movement inside an organization because employees are just as likely to search for alternatives internally. In fact, March and Simon stated that whether there are opportunities within an organization will be related to an employee's desire to leave. Employees may have only vague ideas about external opportunities but can estimate internal availability from "vacancy bulletins, awareness of frequency and mode of filling jobs, and extent of contact with supervisors" (Smith, 1979: 365).

Perceived opportunity for intraorganizational mobility has implications for work motivation and attitudes. Lawler (1971) argued that employees will only work hard to get promoted if they perceive that appropriate positions are available and that promotion decisions are based on work performance. Studies have found that employees who perceive few opportunities for advancement have negative attitudes towards work and their organizations (Kanter, 1977; Kipnis, 1964; Larson, 1982).

The twofold purpose of this study was (1) to investigate the determinants of perceived ease of movement, and (2) to examine the relationship between perceived ease of movement and attitudes towards organizations. Two groups of clerical employees were studied. We chose to study the careers of clerical employees because the recent revolution in office technology and current shortages of qualified clerical help in some areas of the country (Doudna, 1981) have spawned a debate over the careers of office personnel.

Traditionally, the context of clerical jobs has been viewed as "an organized device to get work done with a minimum of disruption" (Cassell, Doctors, & Director, 1982: 8-9) rather than as an opportunity structure. However, Kanter (1977) found that career opportunity at all hierarchical levels can account for the way people involve themselves in their work. Dramatic increases in opportunity may heighten career aspirations, work commitment, and a sense of organizational responsibility for the rank-and-file, as well as for management. Career paths, patterns, and expectations may be different at low and high levels of a hierarchy, but lack of opportunity may lead to the same outcomes. This study was an exploratory attempt to look at the relationships between career expectations, experiences, and attitudes.

HYPOTHESES

Predictors of Perceived Ease of Movement

Three previous models offer suggestions for determinants of perceived ease of movement within an organization. These are Vardi's (1980) and Anderson, Milkovich, and Tsui's (1981) models of intraorganizational mobility, and March and Simon's (1958) model of determinants of perceived ease of movement outside an organization. All three models include personal characteristics as factors influencing expectations about mobility either directly or indirectly. Personal characteristics can be related to perceived ease of movement in two ways. First, individuals have profiles including characteristics such as gender, age, and socioeconomic status that partially determine their attitudes towards work (Vardi, 1978). Second, "such attributes as age, seniority and skill level serve as criteria for mobility when decisions to transfer, promote or demote employees are made by management" (Vardi, 1978: 48). We chose to examine two characteristics that contribute to an individual's profile: age and job tenure. Previous research has found negative relationships between age and expectations of internal mobility (Moore, Miller, & Fossum, 1974; White, 1974). Three possible explanations have been given for this finding. First, age may be an organizational criterion for mobility; second, as people grow older they may become more realistic about the probability of attaining career goals; and third, people may become more cynical as they grow older (Moore, Miller, & Fossum, 1974).

Research on job tenure and plateauing among managers has shown that employees who remain in the same jobs for a number of years are less satisfied with their careers than those who are more mobile (Veiga, 1981). As job tenure increases, unrealistic expectations of rapid promotions are replaced with more accurate perceptions of career opportunities based on experience in searching for better positions. The longer people are unsuccessful in advancing within their organizations, the lower their expectations about career mobility will be.

Hypothesis 1: Age and job tenure will be negatively related to perceived ease of movement.

The match between perceived organizational criteria for mobility and personal characteristics should also influence perceived ease of movement. Anderson and colleagues' (1981) model of intraorganizational mobility included no direct link between individual characteristics and either actual or expected mobility. Instead, the match between individual characteristics and criteria for mobility explained both actual experiences of mobility and individuals' expectations for future mobility. Past research has inferred such a relationship but has not examined it directly.

In this study the variable of interest was perceived organizational criteria for mobility rather than actual criteria, although the two should be correlated. The match between actual criteria and individual characteristics will influence actual mobility, but it is the match between perceived criteria for mobility and individual characteristics that will influence perceived ease of movement. Actual criteria for mobility can be grouped into two categories: ability and nonability (Beehr, Taber, & Walsh, 1980). Nonability factors such as age, race, and gender are not bona fide occupational qualifications for most positions, but employees may perceive them as criteria for promotion. In particular, employees who have not been promoted may have such perceptions, since people tend to attribute failure to factors beyond their control (Beehr et al.). Also, studies have shown that women, black people, and older workers experience less career progress than men, white people, and younger workers (Elliott, 1966; Garbin & Bellweg, 1965; Grusky, 1966; Mahoney & Milkovich, 1982; Raudabough, 1974; Rosenbaum, 1979; Vardi, 1978).

Ability factors such as education, skills, and performance are usually bona fide occupational qualifications on the basis of which management largely makes or at least justifies decisions about promotions. Previous research has shown education to be a predictor of actual mobility (Dalton, 1951; Grusky, 1966; Raudabough, 1974; Rosenbaum, 1979; Sorenson & Tuma, 1981; Stone, 1953). Studies have also found education to be positively related to career aspirations (Hall, 1976; Moore et al., 1974; White, 1974). No one has empirically examined the relationships between employees' perceptions of their own skills and performance and perceived ease of movement. However, it is likely that if employees believe job performance to be a criterion for mobility, those who perceive their performance as above average will perceive themselves as having greater ease of movement than those who do not perceive their performance as above average.

> *Hypothesis 2: The match between perceived organizational criteria for mobility and personal characteristics will be positively related to perceived ease of movement.*

A perceived match between personal characteristics and organizational criteria will be irrelevant if there are no vacancies within an organization or if all vacancies are filled externally. Vardi (1980) suggested that employees will have a greater chance to move ahead when organizations have a policy of filling vacancies from within rather than through an external labor market. Anderson and colleagues (1981) stated that if job vacancies are to be translated into mobility, it must be assumed that those positions are filled from within an organization. We therefore propose that:

> *Hypothesis 3: Perceptions that an organization has a policy of filling vacancies from within will contribute positively to perceived ease of movement.*

In March and Simon's (1958) model, people's visibility to other organizations and the visibility of other organizations to them are factors influencing perceived ease of movement. We argue that visibility is an important factor in perceived intraorganizational movement as well. In particular, we believe that employees' existing jobs can contribute to visibility by allowing them to move around in an organization and discover what jobs are available and by providing opportunities for them to develop skills appropriate for other jobs. Moreover, employees' jobs determine their visibility to those responsible for hiring. Jobs with certain characteristics should thus facilitate the joint visibility of employees and new jobs; these are (1) automony accompanied by responsibility, (2) contact with other people, and (3) task variety.

White (1974) claimed that autonomy is a prerequisite for anyone who plans to stay visible as a job candidate because autonomous employees are held responsible for work outcomes, significant accomplishments can be attributed to them, and they can thus demonstrate potential success in positions of higher responsibility. Using similar reasoning, Kanter (1977) argued against routinized jobs as mobility facilitators.

The amount of contact employees have with organization members outside their own areas should be related to both number of visible job alternatives and their own visibility to those who hire for the jobs. However, organizational policies such as regular postings of job openings in designated areas or listings of vacancies in company newsletters may decrease the need to obtain such information from other people.

Task variety may increase visibility by enabling employees to show broad ranges of accomplishments, demonstrating that they have skills necessary to perform other jobs. As Grandjean and Taylor (1980) argued, "underuse of skills and abilities prevents workers from demonstrating readiness for more responsible, challenging, and renumerative work" (1980: 40).

> *Hypothesis 4: Autonomy, variety, feedback from other people, opportunities to deal with others, and opportunities to learn on a job will be positively related to perceived ease of movement.*

Actual experience of mobility should also influence people's expectations of future mobility (Anderson et al., 1981). However, such a relationship is less likely among lower level employees than among managers because for the first group career ladders are shorter. Thus, they are likely to reach plateaus much more rapidly than managers. For employees who have reached plateaus, actual experiences of mobility may have a positive or a negative relationship to perceived ease of movement, or no relationship; it depends on an individual's position on a career ladder. Although we state no formal hypothesis regarding the relationship between actual past mobility and perceived ease of movement, we will examine it.

Outcomes of Low Perceived Ease of Movement

Several studies have found that actual experiences of mobility are related to attitudes toward jobs and organizations (Bray, Campbell, & Grant, 1974; Grusky, 1966; Martin, 1979; Vardi, 1980; Veiga, 1981). Perceptions of future opportunities for mobility can also be expected to influence job related attitudes, since past mobility is not related to future opportunities for mobility for those employees who have plateaued or reached the tops of their career ladders. If employees who are interested in career advancement do not perceive that opportunities are available within their organizations, they will become dissatisfied and search for alternatives outside. They will compare these alternatives to their present situations and if comparisons favor the alternatives, they will leave (March & Simon, 1958; Mobley, 1977). If comparisons do not favor the alternatives, they may react in several ways. They may continue searching with intentions of quitting as soon as they do find acceptable alternatives. They may become less committed to their organizations. Or they may react to lack of alternatives by lowering their aspirations and desires; previous studies of lower-level employees have found this outcome (Guest, 1954; Kanter, 1977; Vardi, 1980).

> *Hypothesis 5: Perceived ease of movement will be positively related to organization commitment and negatively related to intention to quit for those employees who desire mobility, but not for those employees who report no desire for mobility.*

METHODS

Respondents and Organizations

This study was conducted in the summer in 1982 in two different types of organizations, a university and a state agency, both located in the northeastern United States. The respondent population for the university was drawn from the 1981-82 staff directory. The study included all personnel in the following self-reported job categories: personnel assistant, office assistant, editorial assistant, typist, administrative aide, administrative secretary, secretary, administrative clerk, stenographer, word-processing operator, word-processing supervisor, secretary, and department secretary. Of the 1,176 questionnaires sent out, 300 usable questionnaires were returned for a response rate of 27 percent. The respondents included 6 men and 16 individuals from minority groups. Their average age was 35; their average tenure was eight years; and 83 percent did not have four-year college degrees.

Questionnaires at the state agency were distributed to all clerical staff in the following job categories: stenographers, typists, file clerks, clerks, account clerks, mail and supply clerks, statistical clerks, data entry machine operators, and dictaphone machine transcribers. Of the 1,600 questionnaires distributed, 372 were returned for a response rate of 23 percent. These respondents included 25 men and 20 individuals from minorities. Their average age was 37; their average tenure was eight years; and 92 percent did not have four-year college degrees. The two sets of respondents differed in some respects. Most of the clerical employees at the agency belonged to the Civil Service Employees' Association. Respondents at the state agency were older and had greater tenure than respondents at the university; respondents from the university were more highly educated. We gave data from respondents on race, gender, age, job tenure, education, job title, and job level to personnel representatives at both sites, who indicated that the respondents were representative of their respective populations.

The career development system at the university was employee-driven; individual employees had sole responsibility for initiating transfers. Employees who had been in their current positions at least nine months were eligible to apply for other jobs. Two campus publications posted position openings weekly. Staffing services reviewed all applications and referred qualified employees to the departments that were hiring. The university had a policy that employee candidates were to be given preference over external candidates. The departments that were hiring were responsible for giving employees notice of any action taken, and they conducted interviews and made decisions independently.

The state job-transfer system was also employee-driven but was somewhat more centralized than the university's process because of the civil service exam system. The promotion process included four separate steps: eligibility to compete, competitive examination, canvas, and selection (Center for Women in Government, 1981). Employees could obtain exam announcements and promotion application cards from the agency personnel office.

Candidates who qualified to take exams and who passed got numerical scores and were ranked from highest to lowest on master examination lists. Points were added to scores for seniority. When vacancies occurred, candidates who were on appropriate lists would be asked by the people doing the hiring if they were interested in vacant positions and a selection would be made from the top three candidates who expressed interest. Employees would usually remain on lists for one to two years and then would have to take the exams again. The same exams applied to all agencies; however, each agency had its own list and would consider employees from that list when vacancies occurred. Therefore, for clerical employees, transfers from one agency to another were very unusual.

Procedures

The procedure for distributing the questionnaires differed for the two sites. At the university, the personnel department granted permission to conduct the study with a provision that questionnaires be sent to employees' home addresses and completed on their own time. A cover letter accompanying the questionnaire explained that the study was research for a Ph.D. dissertation on intraorganizational career mobility and the researchers were not acting as consultants. We instructed individuals to return the completed questionnaires directly to us in prestamped envelopes that we provided. At the state agency we conducted the study with the approval of both the civil service and the agency. The personnel department distributed questionnaires and cover letters to supervisors, who distributed the questionnaires to their staff. Employees were given the same explanation for the study and instructed to complete the questionnaires during working hours and return them directly to us in the prepaid envelopes. The agency included an additional cover letter explaining that the purpose of the study was research on career mobility and differentiating it from other studies being conducted at the same time.

Measurement Instruments

Background information on salary scales, career ladders, and internal recruiting and selection procedures was obtained from interviews with personnel staff at both sites. Responses to the questionnaire referred to in the preceding section provided all measures of the variables included in the hypotheses. The questionnaire was eleven pages long. To the extent possible, we used scales that previous studies have validated. Table 2 lists alpha coefficients for all scales.

Job characteristics. Job characteristics were measured by four of Hackman and Oldham's (1980) scales: autonomy, feedback from agents, dealing with others, and skill variety; and one of Rousseau's (1977) scales, opportunity to learn. The solution of a factor analysis of the items supported the a priori scales with one exception: opportunity to learn and skill variety loaded on the same factor. Therefore, we combined these two scales.

Organizational commitment. This was measured using the nine-item, 7-point response scale developed by Mowday, Steers, and Porter (1979). A high score on this scale indicates a high level of commitment.

Intention to quit, perceived organizational policy on filling vacancies, desire for mobility, and perceived ease of movement. These variables were all measured by multi-item scales created specifically for this study. The Appendix gives response formats and items for each scale. Since previous researchers have argued that intention to remain within an organization is one component of commitment (Katz & Kahn, 1966), we performed a factor analysis on the intention-to-quit and commitment items. Results showed two distinct factors. Items for the remaining three scales mentioned above were also factor analyzed; results showed three distinct factors, which supported the a priori categorization of items.

Perceived match. The match between employees' perceptions of organizational criteria for mobility and personal characteristics was measured by comparing responses to items on personal characteristics such as age, gender, and level of job performance with responses to items concerning the importance of specific criteria for mobility. The Appendix includes details on construction of this scale.

Rate of mobility. A work-history table adapted from Vardi and Hammer (1977) was the source for actual information on mobility. We asked employees to record dates, job titles, and job grades for each job they had held since joining their organizations. Frequency and direction of job moves were computed by using job grade information, as explained in the Appendix.

Other measures. Employees were also asked whether they were satisfied with their jobs, whether they would take courses to improve their job opportunities, and what position they wanted to be promoted to if they wanted to be promoted. These were all single items.

ANALYSES AND RESULTS

Career Experiences and Aspirations

Data on career experiences and aspirations provide the context for interpreting the results of the study. Table 1 shows means, medians, standard deviations, and t-values for different mobility measures for the two sites. Results show that rates of overall and upward mobility were greater at the university than at the state agency. Lateral moves were more common at the state agency. Thirty percent of the respondents at the university and 37 percent of the respondents at the agency were still in their first jobs. These mobility rates seem low compared to the mobility rates for blue-collar workers of 27 to 73 percent, depending on the company, found by Vardi and Hammer (1977). However, low average job tenure at both sites could partially explain this discrepancy.

Results also showed that clerical employees were very interested in career advancement. Only 24 percent of university respondents and 18 percent of state agency respondents reported no interest in moving to better

TABLE 1
Comparisons of Actual Mobility for Both Groups

	University			State Agency			
Measures	**Medians**	**Means**	**s.d.**	**Medians**	**Means**	**s.d.**	**Values of *t***
Total mobility	.20	.26	.29	.11	.19	.25	3.26*
Total mobility minus demotions	.19	.25	.29	.10	.18	.23	3.32*
Lateral mobility	0	.20	.11	0	.07	.16	−2.94*
Upward mobility	.17	.05	.27	0	.11	.18	5.48*
Demotions	0	.01	.06	0	.01	.07	− .34

* $p < .05$

jobs. Most respondents—96 percent in the university and 92 percent in the agency—said they would be willing to take courses to improve their career opportunities. Employees did not seek mobility because of dissatisfaction with their current jobs; 71 percent of university respondents and 58 percent of state agency respondents said they liked their jobs.

Tests of Hypotheses 1 through 4

Table 2 shows descriptive statistics and the intercorrelation matrix for all variables for both sites. We computed t-tests to examine differences between the two sites. Results showed that agency respondents reported less task variety, less autonomy, less interaction with others, greater desire for mobility, and greater intention to quit than respondents at the university.

Table 2 also shows that results from both groups supported Hypotheses 1, 2, and 3. Age and job tenure were negatively related to perceived ease of movement. Two variables, (1) the match between perceived criteria for mobility and personal characteristics, and (2) perceived organizational policy of filling vacancies from within, were positively related to perceived ease of movement. Results only partially supported Hypothesis 4. Feedback was positively related to perceived ease of movement for both groups. However, none of the other job characteristics were related to perceived ease of movement for the agency respondents. For the university respondents, autonomy and variety were positively related to perceived ease of movement.

In addition, Table 2 shows that some of the variables hypothesized to be correlated with perceived ease of movement were correlated with each other. Age was positively correlated with job tenure and negatively correlated with the match between perceived organizational criteria for mobility and personal characteristics. Job tenure was also negatively correlated with the match variable. Perceived organizational policy of filling vacancies from within was positively correlated with feedback, autonomy, and variety.

In order to determine the unique contribution of each variable in explaining the variance in perceived ease of movement, we used multiple regression analysis. Independent variables were age, job tenure, perceived organizational policy of filling vacancies from within, match between perceived

TABLE 2
Descriptive Statistics and Intercorrelation Matrix[a]

Variables	Means	s.d.	1	2	3	4	5	6	7	8	9	10	11	12	13
(a) University															
1. Perceived ease of movement[b]	.97	.88	(.83)[a]	−.36	−.30	.21	.34	.27	.00	.14	.21	−.09	.19	−.16	.17
2. Age	36.69	11.43		—	.45	−.09	−.26	.00	.03	.02	.04	−.17	.25	−.21	−.28
3. Job tenure	3.74	4.23				−.10	−.20	.02	.01	−.04	−.01	−.15	.12	−.15	−.30
4. Perceived policy on vacancies	4.76	1.48				(.75)	.07	.25	.04	.16	.18	−.15	.27	−.19	.07
5. Perceived match	1.76	2.21					—	.09	−.04	.07	.03	.30	−.01	.08	.16
6. Feedback	4.81	1.63						(.88)	.19	.41	.52	−.30	.31	−.39	.05
7. Dealing with others	5.66	1.14							(.64)	.27	.43	.01	.03	−.12	.10
8. Autonomy	5.34	1.37								(.79)	.50	−.25	.18	−.37	.11
9. Variety	4.86	1.38									(.88)	−.30	.32	.41	−.12
10. Desire for mobility	4.98	1.39										(.71)	−.32	.44	−.06
11. Commitment	4.53	1.22											(.88)	−.54	.02
12. Intention to quit	2.20	1.54												(.77)	−.09
13. Rate of mobility	.25	.27													

TABLE 2 (continued)

Variables	Means	s.d.	1	2	3	4	5	6	7	8	9	10	11	12	13
(b) State Agency															
1. Perceived ease of movement[b]	.83	.81	(.74)	−.34	−.39	.20	.27	.26	.01	.01	.08	.03	.16	−.13	.06
2. Age	33.90	12.65	—	—	.41	−.06	−.25	.05	.11	.06	−.06	−.18	.27	−.23	−.06
3. Job tenure	4.26	4.58			—	−.13	−.27	−.12	.03	.10	−.04	−.18	.00	−.05	−.32
4. Perceived policy on vacancies	4.68	1.38				(.52)	.16	.18	.05	.12	.19	−.12	.23	−.24	.14
5. Perceived match	1.19	1.95					—	.14	.06	.02	.10	.13	.11	.00	.10
6. Feedback	4.43	1.67						(.87)	.24	.35	.49	−.15	.32	−.30	−.01
7. Dealing with others	5.09	1.29							(.59)	.30	.52	.03	.16	−.28	.03
8. Autonomy	4.77	1.45								(.73)	.58	−.19	.27	−.26	.00
9. Variety	4.06	1.37									(.85)	−.15	.36	−.37	.08
10. Desire for mobility	5.30	1.36										(.70)	−.21	.39	.05
11. Commitment	3.99	1.30											(.88)	−.54	.01
12. Intention to quit	2.99	1.61												(.73)	−.11
13. Rate of mobility	.18	.22													

[a] Correlations greater than .12 are significant at $p < .05$; numbers in parentheses are alpha coefficients.

[b] Standardized.

TABLE 3

Results of Regression of Perceived Ease of Movement on Age, Job Tenure, Perceived Policy on Vacancies, Match, and Four Job Characteristics

	University		State Agency	
Variables	Betas	*F*	Betas	*F*
Rate of mobility	.02	0.16	−.07	2.02*
Age	−.22	12.57*	−.16	8.57*
Job tenure	−.12	3.77*	−.30	28.36*
Perceived policy on vacancies	.08	1.89	.11	4.80*
Perceived match	.21	14.14*	−.14	7.04*
Feedback	.17	7.49*	.18	9.96*
Dealing with others	−.08	1.89	.04	.50
Autonomy	.01	0.02	.00	.00
Variety	.14	4.09*	−.06	.65
R^2	.28		.27	
Adjusted R^2	.25		.25	

* $p < .05$

organizational criteria for mobility and personal characteristics, the four job characteristics, and rate of mobility. Table 3 shows the results of the analyses for both groups. Controlling for the other variables, we found that age and job tenure were negatively associated, and feedback and the match between perceived organizational criteria for mobility and personal characteristics were positively associated with perceived ease of movement for both sites.[1] For the university respondents, variety was positively related to perceived ease of movement. For the agency respondents, perceived organizational policy of filling vacancies from within was positively related to perceived ease of movement. For the state agency respondents, rate of mobility also was positively associated with perceived ease of movement in the regression analysis; however, the zero-order correlation between the two variables was not significant. For the university respondents, rate of mobility was not associated with perceived ease of movement once other variables were controlled.

For both groups of respondents the adjusted R^2 was .25. To determine whether results were population specific, we used a Chow test (Kennedy, 1979) to examine whether the overall equations for the two sites differed significantly. We also compared specific parameters for the two equations by running a regression for data from the two sites combined, including site as a dummy variable and eight interaction terms—site times each of the other independent variables (Cohen & Cohen, 1975).[2] There were no significant differences between the overall R^2s or any of the parameters except for variety on the job. Variety on the job contributed positively to perceived ease of movement for the university respondents, but not for the agency respondents. With the two groups combined, neither rate of mobility nor the interaction of rate of mobility with site were significant. Therefore, the negative relationship between rate of mobility and perceived ease of movement for the agency respondents may only be a statistical artifact.

---

[1] Since age was highly correlated with perceived ease of movement, and age was included in the perceived match scale, there was some concern that the effects of age alone were accounting for the relationship between the perceived match and perceived ease of movement. Thus, we regressed perceived ease of movement on the independent variables, with age removed from the perceived match scale. The results were identical to the first results.

Test of Hypothesis 5

It was hypothesized that desire for mobility would moderate the relationship between perceived ease of movement and both organizational commitment and intention to quit. Following Arnold (1982), we used subgroup analysis rather than moderated regression to test this hypothesis. Employees were divided into two groups: those who desired mobility, as indicated by scores above the median on the scale measuring desire for mobility, and those who were not very interested in moving from their present position, as indicated by scores below the median on the scale measuring desire for mobility. We then used Fisher's Z transformation to compare correlation coefficients between the two subgroups (Arnold 1982). Table 4 shows results. Perceived ease of movement was negatively related to intention to quit and positively related to organizational commitment only for those employees who scored high on the scale measuring desire for mobility. However, for the university respondents, the difference between subgroups for intention to quit was not significant. Therefore, Hypothesis 5 received only partial support.

TABLE 4
Subgroup Analysis of Perceived Ease of Movement with Organizational Commitment and Intention to Quit

Variables	Great Desire for Mobility[a]	z'	Less Desire for Mobility[b]	z'	z
(a) University					
Organizational commitment	.24*	.245	.03	.030	1.69*
Intention to quit	−.14*	.141	−.09	.090	.41
(b) State Agency					
Organizational commitment	−.27*	.277	.09	.090	1.76*
Intention to quit	−.22*	.224	−.04	.040	1.74*

[a] $N = 137$
[b] $N = 128$
[c] $N = 175$
[d] $N = 184$
* $p < .05$

---

[2] Cohen and Cohen (1975) made a distinction between using regression equations for prediction versus explanation. When researchers are interested in using an equation for prediction, they must cross-validate results. When the purpose of a study is explanation, as in this case, the significance of the differences between betas should be tested.

Table 4 shows that several of the variables hypothesized to be determinants of perceived ease of movement were related to commitment and intention to quit. To investigate whether the relationship between perceived ease of movement and each of these two variables was spurious, we regressed both intention to quit and organizational commitment on perceived ease of movement and its hypothesized determinants for those employees who desired mobility. Table 5 shows the results. For the university respondents, perceived ease of movement explained 2 percent of the variance in commitment with other variables controlled; however, perceived ease of movement was not significantly related to intention to quit. For the state agency respondents, perceived ease of movement explained 2 percent of the variance in commitment and 5 percent of the variance in intention to quit, with all other variables controlled.

DISCUSSION

This study showed that young clerical employees who had short job tenure, felt that their personal characteristics matched organizational criteria for mobility, and received feedback from their supervisors and others about how well they were doing on their jobs perceived the greatest opportunities for mobility within their organizations among the employees studied. The relationship between age and perceived ease of movement reflected reality at the university, which had a salary structure of overlapping pay scales for different jobs. Older employees with long tenure were usually at the top of their pay scales, so if they were promoted they would have to be paid at the middle or top of the scale for the next highest grade. Thus, departments often hired younger or new employees whom they could pay at the low ends of scales.

The lack of relationship between perceived organizational policy on vacancies and perceived ease of movement was surprising. However, the respondents who perceived that their organizations filled vacancies externally were less committed to their organizations, and the university respondents having this perception also reported greater intentions to quit. Work attitudes may determine perceptions of policy rather than vice versa. Studies have shown that individuals tend to attribute failure to factors beyond their control but attribute success to their own efforts (Beehr et al., 1980). Therefore, if employees are dissatisfied with their own careers and have been unsuccessfully attempting to transfer to better jobs, they may conclude that their organization prefers to fill positions through the external labor market.

The positive relationship of perceived ease of movement and the match between perceptions of organizational criteria for mobility and personal characteristics provides some evidence that employees assess their opportunities for mobility on the basis of whether they meet the criteria they perceive to be necessary for promotion; however, the amount of variance explained by the scale was not large. More research needs to be conducted on employees' perceptions of criteria for promotion and what factors influence those perceptions.

TABLE 5
Results of Regression of Outcome Variables on Perceived Ease of Movement and Determinants of Perceived Ease of Movement

	University				State Agency			
	Commitment		Intention to Quit		Commitment		Intention to Quit	
Variables	Betas	*F*	Betas	*F*	Betas	*F*	Betas	*F*
Rate of mobility	.01	0.51	−.16	3.83*	−.02	0.08	−.16	4.85*
Age	.23	5.66*	−.30	11.85*	.15	3.90*	−.17	5.28*
Job tenure	−.02	0.54	−.06	0.50	−.13	2.48*	−.02	0.08
Perceived policy on vacancies	.25	7.50*	−.18	5.00*	.18	6.22*	−.07	0.92
Perceived match	−.06	0.33	.09	1.00	.03	0.12	.08	1.18
Feedback	.09	0.71	−.23	5.65*	.16	3.62*	−.15	3.42*
Dealing with others	−.10	1.17	−.07	0.62	.01	0.02	−.07	0.56
Autonomy	.07	0.55	−.28	9.84*	.02	0.06	−.02	0.03
Variables	.16	0.09	−.03	0.07	.26	5.90*	−.24	4.93*
Perceived ease of movement	.19	3.62*	−.06	0.49	.16	3.60*	−.24	9.03*
R^2	.21		.35		.29		.30	
Adjusted R^2	.14		.29		.24		.25	

* $p < .05$

The positive relationship between feedback and perceived ease of movement emphasizes the role supervisors play in the development and career satisfaction of subordinates, particularly in organizations where job transfer systems are decentralized. In this study, positive feedback was not differentiated from negative feedback, but the positive relationship between feedback and the match between perceived organizational criteria for mobility and personal characteristics indicates that employees tend to think of feedback in positive terms only.

The other job characteristics that we suggested would determine visibility of individuals to decision makers and visibility of jobs to individuals were not related to perceived ease of movement. Gould and Penley (1984) found that clerks do not report using the career strategies of networking and creating opportunities as frequently as do managers. We do not know whether their jobs constrain them from using these strategies. In an organization such as the state agency, where mobility is practically nonexistent and likely to be lateral rather than vertical, no use of any strategy is likely to lead to success, nor are the presence of certain job characteristics likely to increase opportunity. More research on the relationship between career strategies and job characteristics for all occupational groups is needed.

The stability of results across the two sites lends credence to these conclusions. However, only 25 percent of the variance in perceived ease of movement was explained. This study limited the hypothesized determinants of perceived ease of movement to variables on an individual level. Future studies should also examine variables at environmental and organizational levels. Environmental variables would include economic conditions, current levels of business or service activity, and demand for particular occupational groups. Organizational variables would include size, technology, number of hierarchical levels, staffing procedures, job posting, career counseling, and career pathing. Anderson and colleagues (1981) hypothesized that "the characteristics of the environment, organization, and work force will have an indirect effect on mobility through their primary influence on the opportunities or vacancies in the organization" (1981: 532).[3] However, whether individuals consider these factors in assessing their career opportunities is open to debate. This study was conducted in a time of high unemployment in towns that were almost entirely dominated by single organizations. Given these environmental characteristics, employees were overly optimistic about intraorganizational opportunities for mobility. One possible explanation for these results is that employees who perceived no opportunities saw no point in participating in the study. The possibility also exists that other occupational groups, such as professionals or upper-level managers, may be more attuned to the influence of environmental and organizational characteristics on career opportunities than lower-level employees.

The study also showed that employees who desired mobility but perceived little opportunity to move to other positions were less committed to their organizations and reported greater intentions to quit than employees who did perceive themselves as having opportunities. However, controlling for other variables, we found that, for the university respondents, perceived ease of movement was not related to intention to quit. Outside job opportunities were probably greater for state agency employees than for the university employees, since the state agency was located in a larger community. Also, university employees were more satisfied with their current jobs, which had more autonomy, variety, and feedback, than employees at the state agency, and many of them had already been promoted.

---

[3] For a review of pertinent literature, see Anderson and colleagues (1981).

These results could have implications for organizations that have career mobility systems and procedures similar to the organizations studied. First, these organizations may want to identify those lower-level employees who desire advancement. Second, if these employees are high performers, organizations could encourage their career development and create ladders for them that permit advancement. If advancement is not possible, jobs could be designed to maximize variety, autonomy, and feedback. Otherwise, in a favorable job market, these organizations could lose their best support staff.

Finally, there is need for research that investigates the meaning of career advancement to various groups of employees and the relationships between actual opportunities for mobility, perceptions of mobility, and performance in organizations. To date, most studies have focused on either aggregate or individual levels of analysis. In a recent study, Scholl (1983) investigated the relationship between objective career-based variables (aggregate level) and turnover and intention to remain (individual level). More studies of this type are needed—but two missing links should be added: employees' interpretations of mobility, and their expectations for the future.

REFERENCES

Anderson, J. C., Milkovich, G. T., & Tsui, A. 1981. A model of intraorganizational mobility. ***Academy of Management Review***, 6: 592–538.

Arnold, H. J. 1982. Moderator variables: A clarification of conceptual, analytical, and psychometric issues. ***Organizational Behavior and Human Performance***, 29: 143–174.

Beehr, T. A., Taber, T. D., & Walsh, J. T. 1980. Perceived mobility channels: Criteria for intraorganizational job mobility. ***Organizational Behavior and Human Performance***, 26: 250–264.

Bray, D. W., Campbell, R., & Grant, D. 1974. ***Formative years in business: A long term AT&T study of managerial lives.*** New York: John Wiley & Sons.

Cassell, F., Doctors, S., & Director, M., 1982. ***A three-company study of the intrafirm mobility of blue-collar and lower-level white-collar workers.*** Final report, Graduate School of Management, Northwestern University, Evanston, Ill.

Center for Women in Government. 1981. ***Barriers to advancement.*** State University of New York at Albany.

Cohen, J., & Cohen, P. 1975. ***Applied multiple regression/correlation analysis for the behavioral sciences.*** New York: John Wiley & Sons.

Dalton, M. 1951. Informal factors in career achievement. ***American Journal of Sociology***, 56: 407–415.

Doudna, C. 1981. Male secretaries: New men of letters. ***New York Times Magazine***, 130 (April 5):121.

Elliott, C. 1966. Age and internal labor mobility of semi-skilled workers. ***Occupational Psychology***, 40: 277–236.

Garbin, A. P., & Bellweg, J. A. 1965. Intra-plant mobility of negro and white workers. ***American Journal of Sociology***, 71: 315–319.

Gould, S., & Penley, L. E., 1984. Career strategies and salary progression: A study of their relationships in a municipal bureaucracy. ***Organizational Behavior and Human Performance***, 34: 244–265.

Grandjean, B. D., & Taylor, P. A. 1980. Job satisfaction among female clerical workers. ***Sociology of Work and Occupations***, 7: 33–53.

Grusky, O. 1966. Career mobility and organization commitment. ***Administrative Science Quarterly***, 10: 488–503.

Guest, R. 1954. Work careers and aspirations of automobile workers. ***American Sociological Review***, 19: 155–163.

Hackman, J. R., & Oldham, G. 1980. ***Work redesign***. Reading, Mass.: Addison-Wesley.

Hall, D. 1976. ***Careers in organizations***. Santa Monica, Calif.: Goodyear.

Kanter, R. M. 1977. ***Men and women of the corporation***. New York: Basic Books.

Katz, D., & Kahn, R. L. 1966. ***The social psychology of organizations***. New York: John Wiley & Sons.

Kennedy, P. 1979. ***A guide to econometrics***. Cambridge, Mass.: MIT Press.

Kipnis, D. 1964. Mobility expectations and attitudes towards industrial structure. ***Human Relations***, 17: 57–71.

Larson, E. 1982. ***Employee commitment to an organization and the effects of perceived ease of movement***. Unpublished doctoral dissertation, State University of New York at Buffalo.

Lawler, E. E., III. 1971. ***Pay and organizational effectiveness: A psychological view***. New York: McGraw-Hill Book Co.

Mahoney, T. A., & Milkovich, G. T. 1972. Internal labor markets: An empirical investigation. ***Academy of Management Proceedings***: 203–207.

March, J. G., & Simon, H. A. 1958. ***Organizations***. New York: John Wiley & Sons.

Martin, T., Jr. 1979. A contextual model of employee turnover intentions. ***Academy of Management Journal***, 22: 313–324.

Milkovich, G. T., Anderson, J. C., & Greenhalgh, L. 1976. Organizational careers: Environmental, organizational and individual determinants. In L. Dyer (Ed.), ***Careers in organizations***: 17–30. Ithaca, N.Y.: New York State School of Industrial and Labor Relations.

Mobley, W. H. 1977. Intermediate linkages in the relationship between job satisfaction and employee turnover. ***Journal of Applied Psychology***, 62: 237–240.

Moore, M. L., Miller, E., & Fossum, J. 1974. Predictors of managerial career expectations. ***Journal of Applied Psychology***, 51: 90–92.

Mowday, R. T., Steers, R. M., & Porter, L. W. 1979. The measurement of organizational commitment. ***Journal of Vocational Behavior***, 14: 224–247.

Raudabaugh, J. 1974. ***Determinants of internal labor market mobility***. Unpublished master's thesis, Cornell University, Ithaca, N.Y.

Rosenbaum, J. E. 1979. Tournament mobility: Career patterns in a corporation. ***Administrative Science Quarterly***, 24: 220–241.

Rousseau, D. M. 1977. Technological differences in job characteristics, employee satisfaction and motivation: A synthesis of job redesign research and socio-technical system theory. ***Organizational Behavior and Human Performance***, 19: 18–42.

Scholl, R. W. 1983. Career lines and employment stability. ***Academy of Management Journal***, 26: 86–103.

Smith, C. 1979. Influence of internal opportunity structure and sex of worker on turnover patterns. ***Administrative Science Quarterly***, 14: 362–381.

Sorenson, A. B., & Tuma, N. B., 1981. Labor market structures and job mobility. In D. Treiman and R. Robinson (Eds.), ***Research in social stratification and mobility: An annual compilation of research***. Greenwich, Conn.: JAI Press.

Stone, R. C., 1953. Factory organization and vertical mobility. ***American Sociological Review***, 18: 20–35.

Vardi, Y. 1978. ***Individual-level and organizational-level determinants of career mobility patterns: An integrative model***. Unpublished doctoral dissertation, Cornell University, Ithaca, N.Y.

Vardi, Y. 1980. Organizational career mobility: An integrative model. ***Academy of Management Review***, 5: 341–356.

Vardi, Y., & Hammer, T. H. 1977. Intraorganizational mobility and career perceptions among rank and file employees in different technologies. ***Academy of Management Journal***, 20: 622–634.

Veiga, J. 1981. Plateaued versus nonplateaued managers: Career patterns, attitudes, and path potential. ***Academy of Management Journal***, 24: 566–578.

White, T. 1974. Production workers and perceptions of intraorganizational mobility. ***Sociological Inquiry***, 44: 121–129.

APPENDIX

The items and response formats for all scales created specifically for this study appear below. Construction of the scale measuring the match between employees' perceptions of organizational criteria for mobility and personal characteristics and computations for actual mobility measures are also discussed in detail.

Unless otherwise noted, we used a 7-point response format, anchored 1 = strongly disagree and 7 = strongly agree for the items in the scales measuring perceived ease of movement, intention to quit, perceived organizational policy or vacancies, and desire for mobility.

Perceived ease of movement

My chances for moving above my present position are high.

It would be easy to find a job in another department.

The following two items were answered according to a 6-point response format, anchored 1 = I have no chances at all, and 6 = I have very high chances:

In your opinion, what are the chances you have to move to another job at the same level as the one you now have, but in a different department?

In your opinion, what are your chances of moving to a job at a higher level within ______?

The four items for this scale were all standardized to correct for differences in response formats and then averaged.

Perceived organizational policy of filling vacancies from within

Job vacancies at ______ are usually filled by people from outside ______.

A ______ employee who applies for another job at ______ has a better chance of getting that job than someone from outside ______ who applies for the job.

Intention to quit

As soon as I can find a better job, I'll leave ______.
I am seriously thinking about quitting my job.
I am actively looking for a job outside ______.

Desire for mobility

I would like a job with more responsibility.
If I'm not promoted from my present job within three to five years, I will be disappointed.
I would feel much better about working at ______ if I were promoted.
I am not interested in moving from my present job. (Reverse scored)

Mobility

To compute rate of mobility, we summed frequencies of lateral and upward mobility and divided the result by organizational tenure. Frequencies of lateral moves, upward moves, and demotions were also computed and divided by organizational tenure.

Perceived match between organizational criteria for mobility and personal characteristics

The perceived match scale was created by comparing two dichotomies: whether respondents perceived each of 16 factors as helping or hindering mobility and whether they had or did not have that particular characteristic. The 16 factors were performing well on the job, being a woman, being a man, having a high skill level, knowing the right people, being Asian, being white, being black, being Indian, being Hispanic, having a great deal of education and training, making a personal effort to get a job, being under 25, being between 16 and 35, being between 36 and 45, and being over 45. Respondents answered these items on a 7-point response format, anchored 1 = greatly hinders and 7 = greatly helps. We then dichotomized responses, with a 1, 2, or 3 categorized as hindering and a 5, 6, or 7 response categorized as helping movement to another job.

These responses were then compared to responses on the following items: gender, race, age, level of education, perceived level of job performance, perceived ability to move to another job, perceived connections with people who could help respondents move to other jobs, and amount of personal effort expended to find a better job. The last four variables were measured on scales with two, three, two, and two items respectively, with Likert-type 7-point response formats. Those who scored less than 4 were categorized as not having that characteristic, while those who scored greater than 4 were categorized as having the characteristic.

A total of 36 comparisons were made. Sixteen of the comparisons were helping conditions. For instance, if being male was perceived as helping mobility, and the respondent was male, the match was scored 1. If being male was perceived as helping mobility, but the respondent was female, the score was 0. Twenty of the comparisons were hindering conditions: if being female was perceived as hindering mobility, and the respondent was female, the match was scored 1. If being female was perceived as hindering mobility, but the respondent was male, there was no match and the score was 0. We summed the hindering and helping conditions separately, and the subtracted the hindering score from the helping score. Therefore, the higher the score the greater was the perceived match between organizational criteria for mobility and the personal characteristics of a respondent.

Jacqueline Landau earned her Ph.D. degree at the New York State School of Industrial and Labor Relations, Cornell University; she is an assistant professor of organizational behavior at the A. B. Freeman School of Business, Tulane University. Her current research interests include absenteeism, careers, and women in management.

Tove H. Hammer earned her Ph.D. degree at the University of Maryland; she is an associate professor of organizational behavior at the New York State School of Industrial and Labor Relations, Cornell University. Her research interests include employee stock ownership, systems of worker participation, and the role of trade unions in labor-management collaboration.

RESEARCH NOTES

© *Academy of Management Journal*
1986, Vol. 29, No. 2, 405–412.

EFFECTS OF AN EXTRINSIC REWARD ON INTRINSIC MOTIVATION: A FIELD EXPERIMENT

PAUL C. JORDAN
University of Alaska-Fairbanks

In their original research, Deci (1971, 1972a, 1972b, 1975) and his associates (Deci, Betly, Kahle, Abrams, & Porac, 1981; Deci, Cascio, & Krussell, 1975) claimed that expectation of contingent rewards undermines intrinsic motivation. There have been more than two dozen studies dealing with the effects of contingent rewards on intrinsic motivation. Unfortunately, lack of a standard terminology has been a major obstacle to integrating these findings. A recent study overcame that problem by standardizing a vocabulary and applying it to a review of the procedures and findings (Ryan, Mims, & Koestner, 1983). The result was a categorization of previous studies that revealed two important gaps in this area of research. First, none of the research compared effects of rewards that were contingent upon performance to effects of rewards that were not contingent upon performance. Second, most previous studies were laboratory ones, and no field research has validated their results (Pinder, 1976).

The purpose of the present study was to test Deci's hypothesis with a better research design than previous studies used. The design of this study incorporated these four features: (1) performance-contingent rewards were compared to rewards that were not contingent upon performance; (2) performance-contingent rewards were offered in combination with regular pay; (3) a field setting was used rather than a laboratory setting; and (4) intrinsic motivation was measured with a scale that has established construct validity. This study also explored possible effects of contingent rewards on satisfaction with pay because previous studies have not examined the effects of expected rewards on people's satisfaction prior to their receiving rewards.

Deci's original work suggested the need to test the following hypotheses:

> *Hypotheses 1: The expectation of a reward that is contingent upon performance will decrease intrinsic motivation.*
>
> *Hypothesis 2: The expectation of a reward that is contingent upon performance will have no effect on satisfaction with pay.*

The author wishes to express his appreciation to Dr. William A. Smith, Jr., Department of Industrial Engineering, and Dr. Donald D. Drewes, Department of Psychology, North Carolina State University, for their assistance in this research.

METHODS

Research Setting

In 1977 the North Carolina General Assembly enacted Senate Bill 722, creating a pilot incentive program for state employees. Pilot programs were established in three units of the state government, including the Western Carolina Center in Morganton, North Carolina, which provides care for 500 mentally retarded and physically handicapped children. The Center's pilot program was tested during a one-year period, from July 1, 1978 through June 30, 1979.

This incentive program promised 225 employees of the Center equal shares of up to 25 percent of all cost reductions at the end of the one-year period. The savings for an administrative unit were to be determined by calculating the difference between the previous two years' average costs and average costs during the experimental year. Because the Center's major expense was labor, the plan's instigators anticipated savings from workforce reductions through attrition and retirement. Analysis of expenditures revealed that a 10 percent decrease in variable expenses was the maximum possible cost reduction. If achieved, this reduction would mean a $500 bonus for each participant in the plan. Participating employees were not told the amount of their potential rewards.

Design

This study used the 2 × 2 × 2 design recommended by Winer (1971) for experiments with repeated measures on a single factor. The design included two treatment conditions. In the first, rewards were contingent upon performance; in the second, rewards were not contingent upon performance. The design also included two work shifts, day and night, and two measurement times, one month before and two-and-a-half months after the beginning of the incentive program. In order to have as many respondents as possible, random assignment of individuals to treatment conditions was not employed.

All potential subjects were tested in their normal work assignments. Preestablished work shifts were preserved so that the study could detect differences arising from variations in supervision.

Subjects

The 48 health care technicians at the Center recruited for the program provided handicapped and mentally retarded children with personal care, including feeding, training, and supervision. Units H and A were selected for participation in the program from four administrative units in the Center because their patients and tasks were similar. Unit H was selected as the group with rewards contingent upon performance because it had the largest number of individuals (32). Unit A, with 16 members, was selected as the group with rewards that were not contingent upon performance. All subjects had at least high school educations, and most were from lower economic strata. No specific training was required to perform their job duties. Table 1 shows biographical data for the subjects.

Measures of Dependent Variables

The study used the Job Diagnostic Survey (JDS) developed by Hackman and Oldham (1975) to measure internal work motivation. Fisher (1978) concluded that this instrument's internal work motivation scale is identical to what Deci (1975) defined as intrinsic motivation.

The Job Descriptive Index (JDI) (Smith, Kendall, & Hulin, 1969) measured satisfaction with pay. No measurement was taken to determine whether or not subjects expected to receive rewards.

Pretest

The JDI and JDS were administered simultaneously to all employees in the two work groups one day prior to the announcement of the incentive program. Supervisors requested employees' participation in the survey but emphasized that it was voluntary and that refusal to participate would not affect their employment. Employees were told that the purpose of the JDI and JDS was to discover how they felt about their jobs, supervisors, co-workers, and working conditions so that management could learn where improvement was needed. In addition, they were told that all responses would be kept confidential and that only group summary scores would be published.

The employees were also asked to complete the biographical data requested on the JDS. Survey responses were grouped by unit and shift, and the names of those who took the survey were recorded on a separate list to ensure that the same individuals would participate in a later survey.

The following day, the Center's administrators announced the state's incentive program for Unit H to all employees. They made no mention of the survey since it was not part of the program. Employees were told that if the pilot program, to begin in one month, was successful it would be implemented in other units later.

TABLE 1
Biographical Data on Subjects

		Age Ranges				
Experimental Groups	**Gender**	**20–29**	**30–39**	**40–49**	**50–59**	**Totals**
Rewards contingent on performance	Men	6	3	1	1	11
	Women	6	5	4	6	21
Rewards not contingent on performance	Men	1	1	0	0	2
	Women	1	5	5	3	14

Posttest

A posttest survey was administered 2½ months into the incentive period and 3½ months after the initial survey. Data on cost reductions were not available at the time of the posttest so subjects still had no idea of the amounts of their potential rewards. Those who had taken the pretest survey were asked to take the posttest survey. Survey responses were again marked and grouped according to unit and shift. Pretest and posttest survey responses were matched by using biographical data such as age group, gender, and educational level; where necessary, matches were based on subjects' handwriting. Random selection of sets of two matched surveys for both conditions followed; after selection the group with rewards contingent upon performance contained 32 individuals and the group with rewards that were not contingent on performance contained 16 individuals.[1]

Analysis

An analysis of variance (Winer, 1971: 540) using the 2 × 2 design referred to earlier in this section was conducted for the measures of intrinsic motivation and satisfaction with pay. If extrinsic rewards affected intrinsic motivation and satisfaction with pay, the interaction of treatment × time (A × C) would be significant. Normality tests were conducted to ensure proper use of parametric statistics and a *t*-test (Winer, 1962) was used to test the simple main effects.

RESULTS

Analysis of variance showed that only the measure for intrinsic motivation had a statistically significant F-ratio, which was 11.486 for the A × C (treatment × time) interaction ($F_{1,44} = 7.25$, $p < .01$). Table 2 shows the means and standard deviations for the two dependent variables.

Figure 1 illustrates two findings. Subjects in both shifts who were in the group with rewards contingent on performance experienced decreases in intrinsic motivation. But in the groups with rewards that were not contingent upon performance, individuals on both shifts experienced increases in intrinsic motivation. Results for satisfaction with pay were not significant. These findings support Hypotheses 1 and 2.

DISCUSSION

The results of this experiment provide empirical evidence that, for these workers in this setting, expectation of rewards contingent upon performance decreased intrinsic motivation. Deci's (1971) theory of cognitive evaluation

---

[1] The pretest was administered to 48 performance contingent and 30 performance non-contingent subjects. Attrition and absenteeism at the time of the posttest reduced these numbers to 42 and 21 respectively. From these matched responses, 32 and 16 were randomly selected to provide for the proportionality of cells frequencies needed for the ANOVA.

TABLE 2
Means and Standard Deviations for Dependent Variables

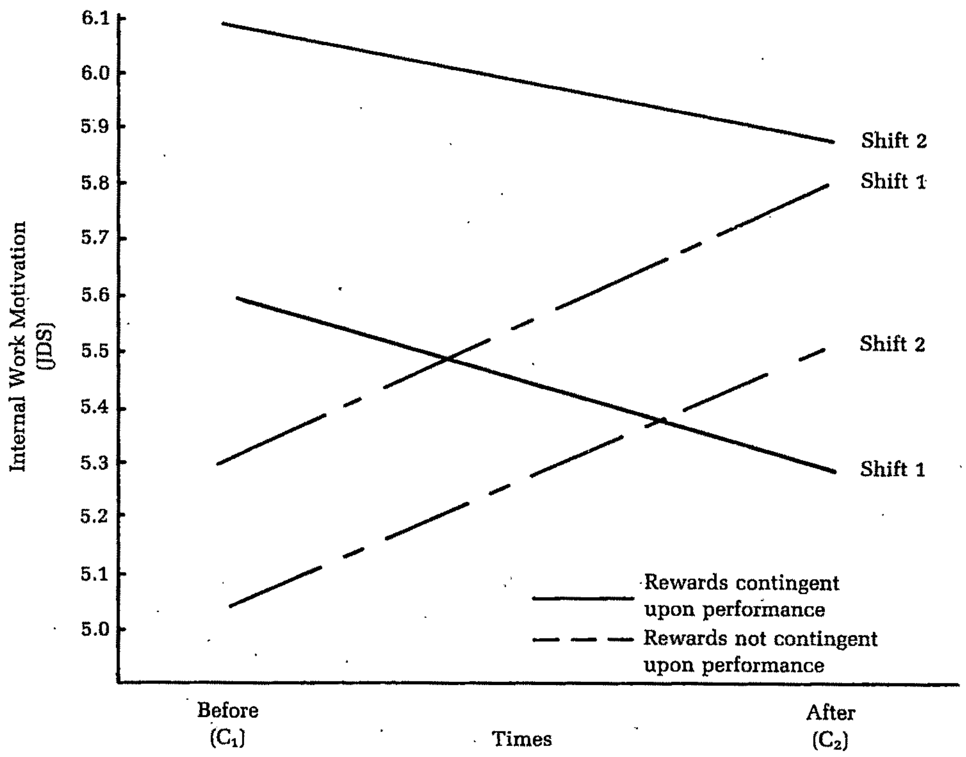
Times of Surveys	Experimental Groups	Shifts	Intrinsic Motivation Means	Intrinsic Motivation s.d.	Satisfaction With Pay Means	Satisfaction With Pay s.d.
1 month before start of incentive program	Rewards not contingent on performance	1	5.30	.350	8.75	8.94
		2	5.04	.704	18.70	67.90
	Rewards contingent on performance	1	5.60	.492	9.37	29.30
		2	6.08	.320	11.10	101.00
2½ months after start of incentive program	Rewards not contingent on performance	1	5.81	.161	6.50	26.70
		2	5.52	.537	14.70	184.00
	Rewards contingent on performance	1	5.29	.879	104.00	121.00
		2	5.89	.277	13.10	111.00

predicted that this would be the case. Generalizing the results of this experiment is inappropriate for two reasons. First, assignment of subjects to treatment conditions was not random. Second, because this was a field experiment, subjects were exposed to a wide variety of uncontrolled stimuli. However, if future research tests the same factors in other field settings and produces the same results, the general applicability of these findings would be strengthened.

Unlike the control subjects in Deci's experiments (1971, 1972a, 1972b), the subjects in the group with rewards that were not contingent upon performance in this experiment were aware of the program offering performance-contingent rewards to their co-workers. DeCharms (1968) provided one possible explanation for the observed increase in intrinsic motivation for the group with rewards that were not contingent upon performance. He theorized that in the absence of contingent rewards, individuals attribute their behavior to internal sources and experience increases in intrinsic motivation.

Organizational researchers may find other aspects of this experiment interesting. For example, the changes in intrinsic motivation occurred in a group incentive plan and in a real workplace. Lawler (1973) stated that shared incentives may negate the bad side effects he observed in individual incentive plans. This study's results indicate that shared incentives do not ameliorate the negative effects of performance-contingent rewards on intrinsic motivations. Pinder (1976) questioned whether Deci's (1971, 1972a, 1972b) results would occur in real organizational settings, where many simultaneously operating reward schedules influence people. Results of this experiment indicate that rewards that are contingent upon performance will

FIGURE 1
Change in Intrinsic Motivation for Two Experimental Groups



decrease intrinsic motivation in the presence of other reward schedules that may be operating.

Although the results reported are statistically significant, the magnitude of the change in scores was relatively small. The Likert-type scale used to measure intrinsic motivation on the JDS has seven gradations. The average change in the different scores was 0.74, or approximately three-quarters of one gradation. This means that change in intrinsic motivation was from a high level to a moderately high level. The absence of any change in attitude toward pay supports DeCharms's (1968) contention that simply contracting to engage in an activity that is contingent upon performance reward is sufficient to undermine instrinsic motivation.

REFERENCES

DeCharms, R. 1968. ***Personal causation: The internal affective determinants of behavior.*** New York: Academic Press.

Deci, E. L. 1971. Effects of externally mediated rewards on intrinsic motivation. ***Journal of Personality and Social Psychology,*** 18: 105–115.

Deci, E. L. 1972a. Intrinsic motivation, extrinsic reinforcement, and equity. ***Journal of Personality and Social Psychology,*** 22: 113–120.

Deci, E. L. 1972b. The effects of contingent and noncontingent rewards and controls on intrinsic motivation. ***Organizational Behavior and Human Performance,*** 8: 217–229.

Deci, E. L. 1975. ***Intrinsic motivation.*** New York: Plenum Press.

Deci, E. L., Betly, G., Kahle, J., Abrams, L., & Porac, J. 1981. When trying to win: Competition and intrinsic motivation. ***Personality and Social Psychology Bulletin,*** 7: 79–83.

Deci, E. L., Cascio, W. R., & Krussell, J. 1975. Cognitive evaluation theory and some comments on the Calder and Staw critique. ***Journal of Personality and Social Psychology,*** 31: 81–85.

Fisher, C. D. 1978. The effects of personal control, competence, and extrinsic reward systems on intrinsic motivation. ***Organizational Behavior and Human Performance.*** 21: 273–288.

Hackman, J. R., & Oldham, G. R. 1975. Development of the Job Diagnostic Survey. ***Journal of Applied Psychology,*** 60: 159–170.

Lawler, E. E. 1973. **Motivation and work organization.** Monterey, Calif.: Brooks/Cole Publishing Co.

Pinder, C. C. 1976. Additivity versus nonadditivity on intrinsic and extrinsic incentives: Implications for work motivation, performance, and attitudes. ***Journal of Applied Psychology,*** 61: 693–700.

Ryan, M. R., Mims, V., & Koestner, R. 1983. Relation of reward contingency and interpersonal context to intrinsic motivation: A review and test using cognitive evaluation theory. ***Journal of Personality and Social Psychology,*** 4: 736–750.

Smith, P., Kendall, L., & Hulin, C. 1969. ***The measurement of satisfaction in work and retirement.*** Chicago: Rand McNally & Co.

Winer, B. J. 1962. ***Statistical principles in experimental design.*** New York: McGraw-Hill Book Co.

Winer, B. J. 1971. ***Statistical principles in experimental design*** (2nd ed.) New York: McGraw-Hill Book Co.

Paul C. Jordan is an associate professor in the Department of Business Administration, School of Management, University of Alaska-Fairbanks. His research interests are in operations management and organizational behavior. He received his Ph.D. degree in industrial engineering from North Carolina State University.

© *Academy of Management Journal*
1986, Vol. 29, No. 2, 412–424.

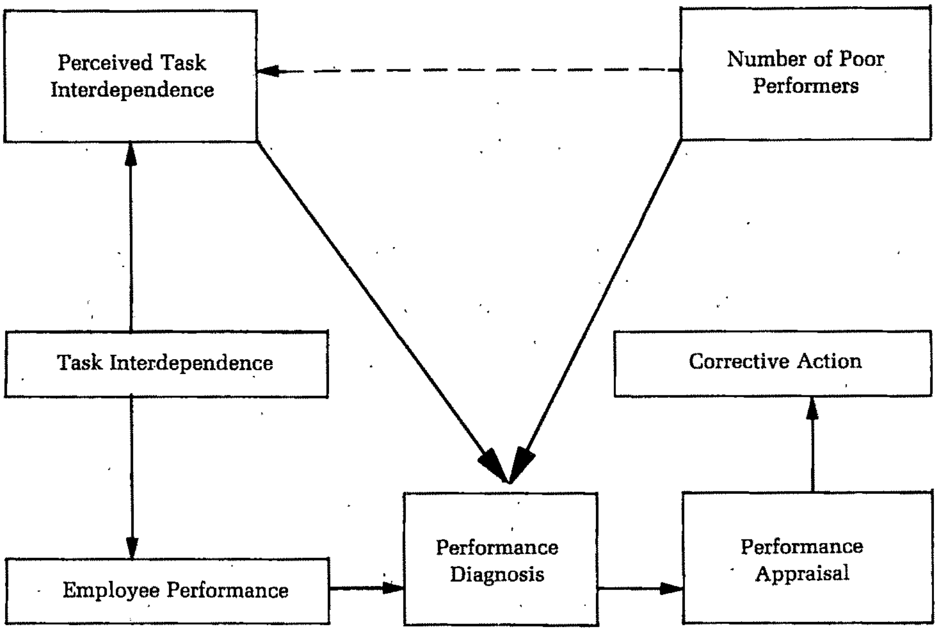
INFLUENCE OF TASK INTERDEPENDENCE AND NUMBER OF POOR PERFORMERS ON DIAGNOSES OF CAUSES OF POOR PERFORMANCE

KAREN A. BROWN
Seattle University
TERENCE R. MITCHELL
University of Washington

Two studies investigated how the number of people performing poorly in a group and the degree of task interdependence among group members influence supervisors when they diagnose the causes of poor performance. We chose to investigate these two factors because of their hypothesized relationships to the appraisal and control of performance. Figure 1 presents a model that serves as a framework for describing the hypothesized influence of perceived interdependence of tasks and number of poor performers on diagnoses concerning performance. The natures of diagnoses influence outcomes such as appraisals of performance and corrective actions. For example, if supervisors diagnose something in a work environment, such as a delay in receiving materials, as causing poor performance, they are likely to direct corrective actions toward the work system; for instance, inventory control might be improved. Additionally, the performance ratings of employees should suffer no adverse effects under such circumstances. On the other hand, if supervisors see employee-centered factors such as motivation or ability as sources of poor performance, they will lower employees' performance ratings and will direct corrective actions toward employees. Improving feedback would be one such action. A research question arises: What influences supervisors to make employee-centered diagnoses, and what influences them to make diagnoses that center on work systems? This research focused on two major variables that we hypothesized to influence the direction of diagnoses of the causes of poor performance.

Support for this research was provided in part by National Science Foundation Grant DAR 79-09792. The authors thank L. R. Beach, T. D. Klastorin, W. T. Newell, and T. G. Schmitt for their comments on the research design. They also wish to thank Bud Borders for his assistance.

FIGURE 1
Performance Control Model



Task Interdependence

A number of authors have studied task interdependence and defined it differently (McCann & Ferry, 1979). However, most authors seem to agree that it can have a significant effect on employees' productivity and the performance of work systems (Cheng, 1983; Dyer, Millen, & Morse, 1978), as well as on employees' morale (Kiggundu, 1983). If task interdependence influences performance, it should be considered in appraisals of performance and in diagnoses of problems with performance (Kane & Lawler, 1979; Peters & Rudolph, 1983; Mitchell & Liden, 1982).

In this research, we expected to find that supervisors make external attributions, diagnosing work environments as causing poor performance, when task interdependence is high. We expected them to make internal attributions, in which they diagnose characteristics of employees as causing poor performance, when task interdependence is low.

Number of Poor Performers

It is useful to consider the extent to which supervisors are sensitive to information about the numbers of people performing poorly in groups and the extent to which they use it appropriately. The more people who are performing poorly, the higher the probability that something common to all employees is to be blamed.[1] Tasks and settings are more likely than employees to be universal forces affecting all members of a group. Thus, we anticipated that supervisors would make internal diagnoses or attributions when only one person performed poorly and external diagnoses when several people performed poorly.

Additional Relationships

Our model also links number of poor performers and perceived task interdependence. Characteristics of work environments provide supervisors with baseline impressions of task interdependence. However, when performance problems occur involving several individuals, supervisors may increase their estimations of the extent to which tasks are interdependent in their groups. Over time, supervisors will continue to modify baseline impressions of task interdependence through observations about the number of people who typically perform poorly.

It is also clear that supervisors will not always make rational decisions as they progress through such diagnostic sequences. For example, Mitchell, Green, and Wood (1981) argued that supervisors have a bias toward finding fault with employees rather than with tasks or environments. Previous research (Jones, 1979) has demonstrated this error in attribution and we expected to obtain similar results. The consequence of a biased diagnosis will be an inappropriate corrective action or an inaccurate performance appraisal, either one of which can detract from future performance levels.

METHODS: LABORATORY STUDY

A laboratory study was designed to provide a controlled test of two variables, task interdependence and number of poor performers. Realistic stimulus materials depicted task interdependence as either high or low and depicted either a group or an individual as poor performers. The study had a 2 × 2 factorial design for the independent variables, with diagnoses of causes of poor performance serving as dependent variables.

Subjects

Subjects were 127 M.B.A. students enrolled in introductory courses on operations management. They provided a relatively homogeneous population whose members had at least some work experience (average years of work experience = 4.6).

---

[1] Kelley's (1967) work on attributions suggested that information about the number of group members performing poorly (called consensus information) is one of a set of performance comparisons that can be made. Comparisons of performance levels for an individual over time (called consistency information) and comparisons of an individual's performance across tasks (called distinctiveness information) also are important. However, Brown (1984) argued that consensus information is the most readily available and accurately assessed type of diagnostic information in work environments.

Procedures

Subjects were provided with a two-page case describing a medium-sized electronics firm. When the subjects finished reading the case, the two independent variables were manipulated through information about a particular job. The result was a 2 × 2 design in which task interdependence and number of poor performers each had two conditions, high and low. We randomly assigned course sections to one of the four treatments.

Next, subjects in each treatment saw a series of slides and heard a tape that described a job assembling circuit boards. These materials had been constructed from information gathered at a major electronics manufacturing firm. Some of the slides showed diagrams of the flows of material and information for the circuit-board-assembly work center, and others were photographs taken on the shopfloor of the electronics firm. In the accompanying audio tape, an employee—actually a research assistant—described his job in the work center. The slides and description of the task were identical in all conditions except for two manipulations. Half of the subjects received information on work flows that indicated high task interdependence and the other half received information indicating low task interdependence. Figure 2 shows how flows of materials and information were depicted for both conditions. As this figure reveals, task interdependence was both *initiated* and *received* because it flowed in two directions (Kiggundu, 1983).

Next, subjects reviewed information about a performance problem. An individual employed in the work center—but not the person who had described the job on the tape—was depicted as having produced at 40 percent below expected levels of output during the previous week. A slide illustrated the decline in performance graphically; then a second slide provided comparative information about other employees in the workgroup. Half of the subjects saw a slide revealing that the focal employee was the only one experiencing a decline in output, and half saw a slide revealing that other group members were experiencing similar declines. This procedure manipulated the variable for number of poor performers. Figure 3 shows the diagrams used in this manipulation.[2]

After their exposure to the two manipulations, subjects received a questionnaire designed to assess their perceptions of the causes of the individual's performance problem. Subjects rated nine items[3] on how likely each was to have caused the individual's performance to decline. Responses were on 5-point scales. Other items provided a manipulation check by requiring subjects to rate the extent to which they felt employees in the work center depended on (1) work materials and (2) information to perform their jobs.

---

[2] As a result, the manipulation actually blended consensus information with what Kelley (1967) termed consistency information. However, we considered this mix of the two types of information necessary for creating realistic performance charts. Also, for the focal employee who was being rated, the consistency information was exactly the same in both conditions.

[3] The nine items used were: To what degree do you feel that something about the employee (e.g., his effort or ability) is likely to have caused the performance problem; that something in the work environment or something about the job (e.g., problems with materials or information) may have caused the performance problem; that the employee's performance problem may have been caused by personal problems; that a delay in the arrival of incoming work materials may have contributed to the employee's problem; that the quality of incoming work materials may

(Continued)

RESULTS: LABORATORY STUDY

The manipulation checks were used to test the strength of the manipulation of perceived task interdependence. The mean of the two scaled items produced an index of perceived task interdependence. Results of a 2 × 2 ANOVA on this index showed a main effect for task interdependence ($F_{1,123} = 10.93$, $p<.001$) and no main effect for number of poor performers. This verified that the task interdependence manipulation had been effective. There was also an interaction effect ($F_{1,123} = 5.55$, $p<.023$). Specifically, subjects revised their perceptions of task interdependence upward when confronted with a performance problem involving several people. This result supported the hypothesized moderating effect of information about number of poor performers on perceived task interdependence.

Testing Hypotheses

A two-way multivariate analysis of variance between subjects was performed on scaled responses to questions about the causes of the decline in performance. The analysis revealed significant effects for task interdependence ($F_{8,112} = 3.197$, $p<.003$); number of poor performers ($F_{8,112} = 117.69$, $p<.001$); and the interaction term ($F_{8,112} = 2.58$, $p<.013$).

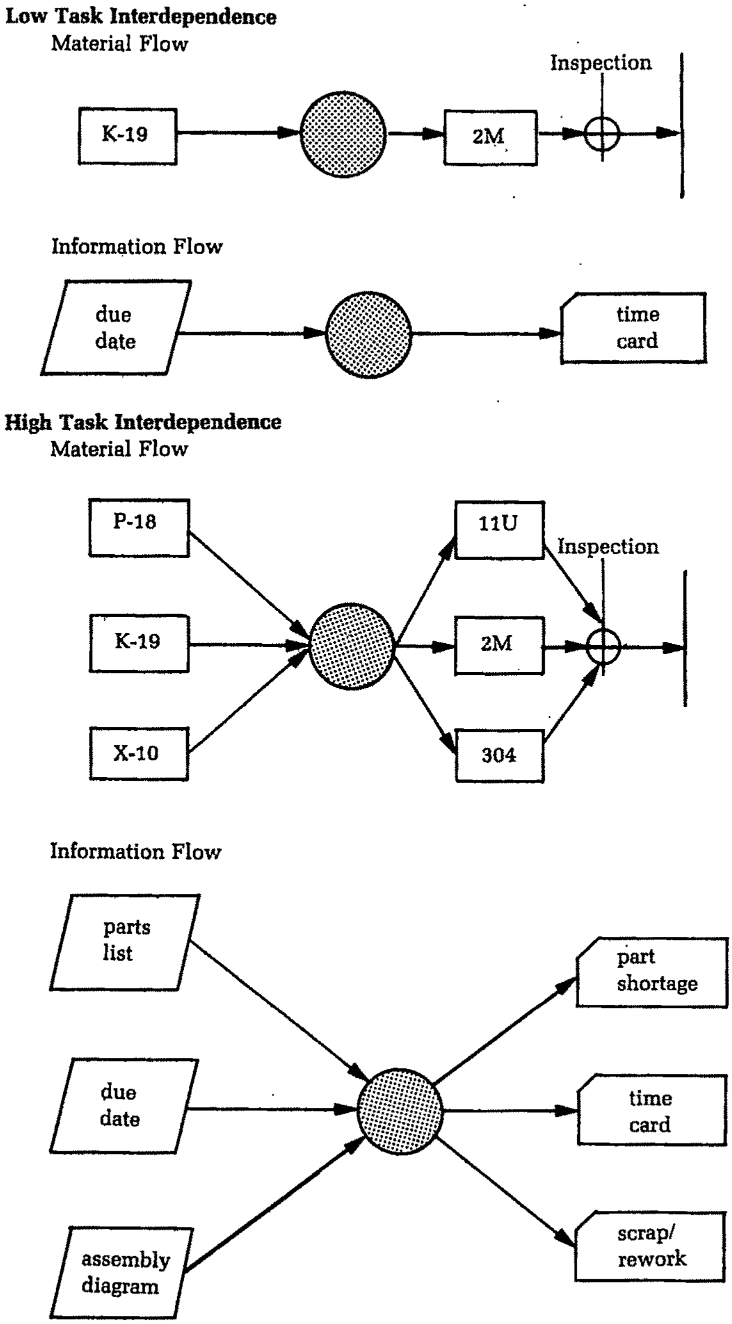
As an aid to the interpretation of the significant MANOVA results, we calculated mean scores on items from the questionnaire to create two indices. One index was based on responses to the two items suggesting employee-centered or internal causes ($r = .38$, $p<.001$). The other was based on responses to the seven items suggesting external causes related to the work system ($\alpha = .776$). The mean scores for these two indices were not significantly different, demonstrating that the anticipated bias toward employee-centered diagnoses did not occur. We next subjected the two indices to a two-way analysis of variance.

The F-statistics for the cell means shown in Table 1 reveal that both number of poor performers ($F_{1,123} = 60.45$, $p<.001$), and task interdependence ($F_{1,123} = 6.34$, $p<.013$) influenced diagnoses that centered on employees. The effects were in the hypothesized direction: the individual employee was more likely to be blamed if there was only one poor performer or if task

---

have contributed to the employee's problem; that problems encountered after completed circuit boards leave the work center (e.g., inspection errors or breakage) may have contributed to the employee's problem; that delays in the arrival of work-related information may have contributed to the employee's problem; that problems with the accuracy or quality of incoming work information may have contributed to the employee's problem; that problems encountered after information leaves the work center (e.g., keypunching errors or lost records) may have contributed to the employee's problem?

FIGURE 2
Task Interdependence Resource Flows: Laboratory Study



Low Task Interdependence
Material Flow

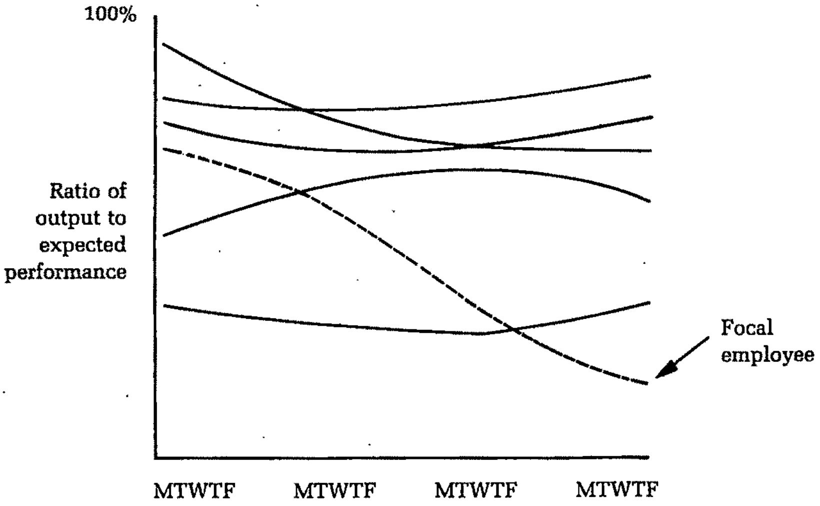
Information Flow

High Task Interdependence
Material Flow

Information Flow

FIGURE 3
Performance Charts For Number of Poor Performers: Laboratory Study

One Poor Performer



Several Poor Performers

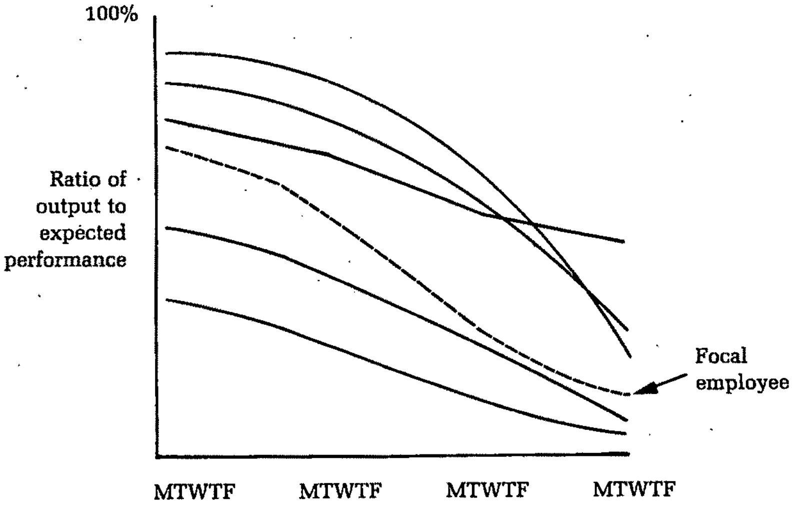


TABLE 1
Mean Ratings for Cells: Laboratory Study

	Task Interdependence[b]		
Number of Poor Performers[a]	**Low**	**High**	**Overall**
(a) Index of Employee-Related Causes (Internal)			
One	3.59 n=33	3.03 n=31	3.32 n=64
Several	2.32 n=37	2.17 n=26	2.27 n=63
Overall	2.92 n=70	2.65 n=57	
(b) Index of System-Centered Causes (External)			
One	2.11 n=33	2.72 n=31	2.41 n=64
Several	3.10 n=37	3.03 n=26	3.07 n=63
Overall	2.63 n=70	2.86 n=57	

[a] $F_{1,123} = 60.45$, $p<.001$ for internal causes; $F_{1,123} = 42.48$, $p<.001$ for external causes.
[b] $F_{1,123} = 6.34$, $p<.013$ for internal causes; $F_{1,123} = 7.22$, $p<.003$ for external causes.

interdependence was low. There was no interaction effect for the employee-centered causes.

Both independent variables also significantly affected system-related diagnoses (number of poor performers, $F_{1,123} = 42.48$, $p<.001$; task interdependence, $F_{1,123} = 7.22$, $p<.008$). These results were also in the hypothesized direction. Causes related to the work system (external attributions) received their highest ratings when there were several poor performers or when task interdependence was high.

In addition to the two main effects, there was a significant interaction effect ($F_{1,123} = 10.34$, $p<.002$) on the index of system-related diagnoses. The interaction term shows that task interdependence influenced attributions of poor performance for the individual more than it influenced attributions of poor performance for the group. This suggests that when several people are performing poorly, information about task interdependence contributes little to diagnoses of the causes of the poor performance because information about the number of poor performers has already pointed to the work system as the cause of the problem.

FIELD STUDY

A field study was conducted to assess the extent to which task interdependence and information about number of poor performers influence supervisors' diagnostic judgments in actual work settings. Additionally, it provided an opportunity for exploring manufacturing supervisors' opinions about the causes of declining output.

Respondents

The field study involved one-hour interviews with supervisors and managers in 12 manufacturing firms. We chose this group from the respondents to a mail survey; the firms represented a variety of industries, including producers of electronics equipment, garments, ships, industrial equipment, wood products, graphics, and industrial plating.

From one to six people were interviewed in each firm, depending on its size and on the availability of personnel for participation in the study. In total, we interviewed 46 people.

Measures

In each firm, a specific semiskilled task was selected as the focus for all responses. During the first part of each interview, respondents rated the degree of task interdependence associated with the focal tasks. They rated interdependence between and within work units with regard to materials, machinery, scheduling information, specialized manpower, and teamwork. All scales had 5-point response formats.

Next were scaled items designed to tap respondents' personal theories regarding the causes of poor performance. Two scenarios were described: one in which an individual was performing a focal task at 20 percent below the expected level of output, and one in which an entire work group was performing the task at 20 percent below the expected level.

Respondents rated eight causal factors as to the likelihood of their having contributed to each of the two performance problems. We adapted these causal factors from Buffa's (1976) description of assignable causes and from Deming's (1975) listing of the common causes of variation in performance. Using a 5-point response format, respondents rated each factor as (1) a cause of the individual's poor performance and (2) a cause of the group's poor performance.

RESULTS: FIELD STUDY

The mean ratings for the eight causal factors were, in order of magnitude: motivation, 3.22; quality of materials, 2.95; training, 2.94; ability, 2.58; delays of materials, 2.58; performance of machines or equipment, 2.51; capacity of machines, 2.50; scheduling, 2.23. Thus, respondents saw motivation as the most likely cause of poor performance. Aggregations of the employee-centered causes (motivation, training, ability) and the causes centered on work systems (delays in materials, quality of materials, performance of machines, capacity of machines, and scheduling) showed that respondents saw the employee-centered factors ($\bar{x} = 2.91$) as significantly more important than the work-system related factors ($\bar{x} = 2.50$; $F_{1,45} = 7.8$; $p < .01$). As predicted, managers and supervisors saw employees, rather than environment, as the cause of performance problems.

Testing Hypotheses

Respondents were separated into two groups on the basis of their mean responses on the ten items measuring interdependence ($\alpha = .743$). We made the split at the median. The aggregated measure used to split the two groups ($\bar{x} = 3.45$ and 4.44) was significantly higher in one group than in the other ($F_{1,44} = 83.97$, $p < .001$). We performed a multivariate analysis of variance on responses to the eight scaled items about the causes of declining output. The model was based on 2 × 2 design in which task interdependence was high or low, and number of poor performers was all or one. Task interdependence was a categorical variable and number of poor performers was a manipulated variable. Multivariate effects for task interdependence were not significant. There were no interaction effects. However, the multivariate result for information about number of poor performers was significant ($F_{9,36} = 8.02$, $p < .001$).

A simplified analysis based on aggregated indices of employee-centered and system-related causes was performed to examine the general effects of the two independent variables. We entered the indexed variables into a 2 × 2 ANOVA model as dependent variables. Table 2 reports means for cells. The index for employee-centered causes ($\alpha = .60$) was influenced significantly by information about the number of poor performers ($F_{1,44} = 48.28$, $p < .001$), but not by task interdependence. Information about the number of poor performers ($F_{1,44} = 18.2$, $p < .001$) influenced the index for system-related causes ($\alpha = .76$), and task interdependence had a near significant effect ($F_{1,44} = 3.6$, $p < .07$). Neither index resulted in an interaction effect.

DISCUSSION

To summarize the results of the two studies, number of poor performers affected diagnoses of reasons for poor performance significantly in both the laboratory and field studies. Task interdependence had its only significant effect in the laboratory study in which graphic displays and job descriptions made its presence clear. The lack of a significant effect for task interdependence in the field has several possible explanations. First, supervisors may not pay enough attention to the variable. Additionally, there may be a defensive element to this lack of acknowledgment. Supervisors may view themselves as responsible for moderating the effects of task interdependence, but avoid including it in diagnoses because doing so would point some blame at them (Jones, 1979).

Also, task interdependence was not manipulated in the field study, and differences between the high and low groups may not have been sufficient to produce differences in diagnoses. The range of interdependence may have been restricted because all respondents focused on semiskilled tasks that had similar characteristics.

It is difficult to determine the true cause of the weakness of the effects for interdependence in the field study. However, the methodology and results

TABLE 2
Mean Ratings for Cells: Field Study

	Perceived Task Interdependence[b]		
Number of Poor Performers[a]	**Low**	**High**	**Overall**
(a) Index of Employee-Related Causes (Internal)			
One	3.25 n=23	3.48 n=22	3.36 n=46
Several	2.41 n=23	2.51 n=23	2.46 n=46
Overall	2.82 n=23*	3.00 n=23*	
(b) Index of System-Related Causes (External)			
One	2.17 n=23	2.31 n=32	2.24 n=46
Several	2.54 n=23	3.00 n=23	2.77 n=46
Overall	2.30 n=23*	2.71 n=23*	

[a] $F_{1,44} = 44.28$, $p<.001$ for internal causes; $F_{1,44} = 18.2$, $p<.001$ for external causes.
[b] Non-significant for internal causes; $F_{1,44} = 3.6$, $p<.07$ for internal causes.
*Number of poor performers was a within-subjects manipulation, so the overall n is the same as each of the cell n's.

of the laboratory study demonstrated that when task interdependence is clearly described and subjects have no responsibility for buffering its effects on employees' performance, they will use it in diagnoses of causes of poor performance.

CONCLUSIONS

This research makes some contributions and suggests a need for further study. We demonstrated both in the laboratory and the field studies that information about the number of poor performers has a significant effect on diagnoses of the causes of poor performance. Previous research has demonstrated these effects in educational settings (MacArthur, 1971), but they have not been demonstrated previously with respect to performance diagnoses in work settings. The interdependence effects that emerged in the laboratory also offer a new contribution in the area of performance diagnosis.

Another contribution comes in the form of a nonsignificant finding. It was a surprise that interdependence did not affect the supervisors' perceptions of the causes of performance problems in the field study. Although we gave some possible post hoc explanations for this finding, we still do not have sufficient evidence to choose among these alternatives. Subsequent work is needed in this area.

In addition to the finding related to the effects of the two information sources, the data from the field study showed a general tendency toward internal attributions. This confirms results from other studies that have demonstrated observers' propensities to blame people rather than environments for performance problems. One possible implication is that improvement in productivity will come through a focus on changing employees' abilities, levels of motivation, attitudes, and so forth. However, previous research provides some evidence that increased awareness of social, task, situational, technological, and environmental limitations on performance is needed (Peters, O'Connor, & Rudolph, 1980).

It also appears that supervisors may be reluctant to use information on task interdependence. Until organizations remedy this problem, perhaps through training, it is likely that diagnoses of the causes of poor performance will continue to focus on individuals and possibly continue to be inaccurate or incomplete. This research highlights this point by demonstrating that only when information on task interdependence was made highly salient in the laboratory study did it have a noticeable effect on diagnosis. Future work needs to focus not only on what information people use in making diagnoses, but also what information they should use.

REFERENCES

Brown, K. A. 1984. Explaining group poor performance: An attributional analysis. ***Academy of Management Review***, 9: 54–63.

Buffa, E. S. 1976. ***Operations management: The management of productive systems.*** Santa Barbara, Calif.: John Wiley & Sons.

Cheng, J.L.C. 1983. Interdependence and coordination in organizations: A role-system analysis. ***Academy of Management Journal***, 26: 156–162.

Deming, W. E. 1975. On some statistical aids toward economic production. ***Interfaces***, 5(4): 1–15.

Dyer, J. S., Millen, R. A., & Morse, J. J. 1978. A framework for the study of work settings. ***Management Science***, 24: 1393–1403.

Jones, E. E., 1979. The rocky road from acts to dispositions. ***American Psychologist***, 34: 107–117.

Kane, J. S., Lawler, E. E. 1979. Performance appraisal effectiveness: Its assessment and determinants. In B. M. Staw (Ed.), ***Research in organizational behavior***, vol. 2: 425–478. Greenwich, Conn.: JAI Press.

Kelley, H. H. 1967. Attribution theory in social psychology. In D. Levine (Ed.), ***Nebraska symposium on motivation***, 15: 192–238. Lincoln: University of Nebraska Press.

Kiggundu, M. N. 1983. Task interdependence and job design: Test of a theory. ***Organizational Behavior and Human Performance***, 31: 145–172.

MacArthur, L. A. 1971. The how and what of why: Some determinants and consequences of causal attributions. ***Journal of Personality and Social Psychology***, 22: 171–193.

McCann, J. E., & Ferry, D. L. 1979. An approach for assessing and managing inter-unit interdependence. ***Academy of Management Review***, 4: 113–119.

Mitchell, T. R., Green, S. G., & Wood, R. E. 1981. An attributional model of leadership and the poor performing subordinate: Development and validation. In B. M. Staw & L. L. Cummings

(Eds.), ***Research in organizational behavior***, vol. 3: 197–234. Greenwich, Conn.: JAI Press.

Mitchell, T. R., & Liden, R. C. 1982. The effects of social context on performance evaluations. ***Organizational Behavior and Human Performance***, 29: 241–256.

Peters, L. H., & Rudolph, C. J. 1983. ***Embedding performance appraisal within an organizational context.*** Paper presented at the annual meetings of the American Psychological Association, Washington, D.C.

Peters, L. H., O'Connor, E. J., & Rudolph, C. J., 1980. The behavioral and affective consequences of performance relevant situational variables. ***Organizational Behavior and Human Performance***, 25: 79–96.

Karen A. Brown received her Ph.D. degree in business administration from the University of Washington. She currently is an assistant professor in the Albers School of Business at Seattle University. Her research interests include work design, productivity, and managerial decision making.

Terence R. Mitchell received his Ph.D. degree from the University of Illinois. He currently is Edward G. Carlson Distinguished Professor of Business at the University of Washington. His research interest include motivation, leadership, and decision making.

© *Academy of Management Journal*
1986, Vol. 29, No. 2, 424–433.

ADDITIVE INFORMATION PROCESSING AND THE RELATIONSHIP BETWEEN EXPECTANCY OF SUCCESS AND MOTIVATIONAL FORCE

ADRIAN HARRELL
University of South Carolina
MICHAEL STAHL
Clemson University

In the force model posited by expectancy theory (Vroom, 1964), the level of motivational force acting on a person is determined by multiplying the model's valence and expectancy variables. Logically, this implies that increases in expectancy result in proportionate increases in motivation. However, some recent research (Kahneman & Tversky, 1979; Rynes & Lawler, 1983) has suggested that increases in expectancy may result in disproportionately small increases in motivation. In order to consider this issue, we employed data from five groups to examine the proposal that increases in the expectancy of success result in declining marginal increases in motivational force.

For more than two decades, expectancy theory (Vroom, 1964) has been one of the more popular theories of motivation (Connolly, 1976; Heneman & Schwab, 1972; House, Shapiro, & Wahba, 1974; Locke, 1975; Mitchell, 1974, 1979; Mitchell & Beach, 1977; Schwab, Olian-Gottlieb, & Heneman, 1979;

Stahl & Harrell, 1981, 1983; Wahba & House, 1974), with most research efforts focusing on the force model. Similarly, the expected utility model has long been popular among economists and researchers concerned with normative decision theory (Keeney & Raiffa, 1976; Samuelson, 1967) for explaining human behavior. A number of researchers have noted the close correspondence between these two models (e. g., Harrell & Stahl, 1984; Mitchell & Beach, 1977; Vroom, 1964: 19). The expectancy variable in the force model and the expectation variable in the expected utility model both refer to the probability that a choice or act will lead to an outcome. Valence and utility both refer to affective orientations toward outcomes (Samuelson, 1967: 417–419; Vroom, 1964: 15). In addition, both models imply that individuals employ multiplicative information processing. The models' premise is that people multiply valence (utility) by expectancy (expectation) to determine levels of motivation (expected utility); this implies that increases in expectancy (expectation) should result in proportionate increases in motivation (expected utility). However, investigators have criticized both the expectancy force model and the expected utility model. Methodological and measurement difficulties experienced in early research on expectancy theory, such as the problems discussed by Kopelman (1977), Schmidt (1973), and Wolf and Connolly (1981), may have resulted in some decline in the force model's popularity. Similarly, the descriptive validity of the expected utility model has been questioned over the years (e. g., Allais, 1953; Coombs, 1975; Ellsberg, 1961), with recent criticism from Kahneman and Tversky (1979). The close correspondence of these models suggests that criticisms of one model logically also apply to the other.

Mitchell and Beach (1977) and Zedeck (1977) suggested that behavioral decision theory could be used to avoid the problems experienced by earlier expectancy theory researchers. Acting on these suggestions, in two previous studies (Harrell & Stahl, 1984; Stahl & Harrell, 1981) we used judgment analysis to examine the premise of multiplicative information processing implicit in the expectancy force model. We concluded that most of the individuals studied employed additive information processing to arrive at their decisions concerning motivation, with a minority employing the multiplicative information processing assumed by the force model. Accordingly, we proposed the force model be modified to acknowledge these individual differences in information processing (Stahl & Harrell, 1981). Rynes and Lawler (1983), replicating and extending our (1981) research, employed several different analysis techniques and also concluded that only a few of their participants employed multiplicative information processing. Butler and Womer (1985), employing an analysis technique different from those that either we or Rynes and Lawler used, also concluded that a minority of their participants employed the multiplicative information processing assumed by the force model. We suggested that our (1981) results implied that individuals with high levels of education in mathematics and statistics appear more likely to employ multiplicative information processing than are others. Butler and Womer (1985) tested this proposal and concluded that people's facility with mathematics could be used to predict their use of either multiplicative or additive information processing to arrive at decisions concerning motivation. Since both the expectancy force model and the expected utility model imply that individuals employ multiplicative information processing, these results logically apply to both models.

Kahneman and Tversky (1979), after presenting criticisms of the expected utility model that also logically apply to the expectancy force model, proposed an extension of expected utility theory called prospect theory. Prospect theory incorporates many of the empirical findings presented in the literature on behavioral decision theory (Slovic, Fishhoff, & Lichtenstein, 1977; Slovic & Lichtenstein, 1971). Prospect theory assumes two phases, an initial editing phase and an evaluation phase. During the editing phase, people eliminate some alternatives and frame acts, outcomes, and contingencies. During the evaluation phase, they evaluate prospective outcomes and make choices. The prospect theory model, which has not yet been empirically tested, is intended to describe how individuals make choices during the evaluation phase. As with the expectancy force model and the expected utility model, the prospect theory model assumes that individuals employ multiplicative information processing when they evaluate prospective outcomes. A principal difference between the prospect theory model and the traditional expected utility model is that prospect theory assumes utility to be multiplied by a weight that is a nonlinear function of expectation, rather than by the actual value of the expectation variable. Kahneman and Tversky presented a hypothetical weighting function (1979: 283) that suggests individuals tend to overweight low probabilities and underweight moderate and high probabilities. They also presented an S-shaped value function that is convex for negative nonzero outcomes and concave for positive nonzero outcomes (1979: 279). Since they hypothesized that the part of the value function relating to positive nonzero outcomes[1] is concave, increases in expectation should result in choices that imply declining marginal increases in expected utility. On the basis of findings presented in their research on expectancy theory, Rynes and Lawler (1983) proposed similar relationships between the expectancy of success and motivational force.

METHODS

This study employed data from five diverse groups to examine the proposal that increases in the expectancy of success result in declining marginal increases in motivation. Although the focus of this research was upon the expectancy force model, the results logically apply to the expected utility model and in some instances to the prospect theory model as well. We previously used the data employed in this study to examine (1) the expectancy force model's premise of multiplicative information processing, and

---

[1] The upper right quadrant of Figure 1 in Tversky and Kahneman (1981: 454) depicts this part of the value function.

(2) the force model's ability to explain the decisions reached by individuals (Harrell & Stahl, 1984; Stahl & Harrell, 1981). Those studies did not consider the relationship between expectancy of success and motivational force, the issue of primary interest here.

Our previous research described how the judgment analysis approach that those researching behavioral decision theory have employed can be used to perform research on expectancy theory. Thus, we present only a brief description of this process. With the judgment analysis approach, each participant must complete a series of decision making cases, in which the elements of each case are designed to correspond to the various elements of the valence and force models of expectancy theory. Factorial designs are incorporated into the information presented in the cases. Initially, each case presents a number of second-level outcomes and their relationships to the first-level outcome of interest. On the basis of this information, each subject reaches a decision as to the attractiveness (valence) of the first-level outcome. This decision provides an operational measure of the valence of the first-level outcome. After the expectancy of success is presented, the subject then reaches an effort-level decision that provides an operational measure of the motivational force acting upon that individual under the particular circumstances of the case. Data are analyzed separately for each individual. Usually, regression analysis is employed for the valence model, with the beta weight associated with each second-level outcome indicating its valence (Stahl & Harrell, 1983). Hierarchical regression analysis (Arnold & Evans, 1979; Cohen, 1978; Cohen & Cohen, 1975; Harrell & Stahl, 1984; Stahl & Harrell, 1981) is usually employed to examine the force model's multiplicative premise.

Five groups of subjects completed four exercises in decision making, with different groups doing different exercises. The exercises provided operational measurements of (1) motivation to get a particular job (groups 1 and 2), (2) motivation to accomplish budgeted cost goals (group 3), and (3) motivation to earn an "A" in an academic course (groups 4 and 5). Our earlier studies provide detailed descriptions of the subjects and the decision making exercises (Harrell & Stahl, 1984; Stahl & Harrell, 1981). The situations presented in all four exercises involved only positive nonzero outcomes and therefore related only to the concave portion of the prospect theory value function (Kahneman & Tversky, 1979: 289; Tversky & Kahneman, 1981: 454).

RESULTS

The results of the data analysis are presented separately for each of the five groups.

In group 1, 59 men who were seniors in accounting at a southwestern university completed a decision making exercise, with each subject's effort-level decision values providing operational measurements of his motivation to get a particular job. Each subject provided eight decision values for each of three expectancy levels (0.00, 0.30, and 0.60). Since a factorial design was incorporated into the exercise, the information upon which the three sets of eight decisions were based differed only as to the expectancy levels involved. Therefore, if increases in expectancy (expectation) result in declining marginal increases in the value of motivational force, the differences between each subject's corresponding decision values when expectancy equaled 0.30 and when expectancy equaled 0.00 will be greater than the differences between each subject's corresponding decision values when expectancy equaled 0.60 and when it equaled 0 30. In this and in subsequent analyses, we used a paired-samples t-test to examine this hypothesis.[2] The results (df = 471, t = 4.6, $p < .0001$) support the proposal that increases in expectancy result in declining marginal increases in motivation.

In group two, 75 junior and senior administrative management students at a southeastern university completed a decision making exercise, with each subject's effort-level decision values providing operational measurements of the person's motivation to get a particular job. Each subject provided eight decision values for each of three expectancy levels (0.05, 0.50, and 0.95). If, as proposed, increases in expectancy result in declining marginal increases in motivational force, the differences between each subject's corresponding decision values when expectancy equaled 0.50 and when it equaled 0.05 will be greater than the differences between each subject's corresponding decision values when expectancy equaled 0.95 and when it equaled 0.50. The results (df = 591, t = 4.4, $p < .0001$) support the proposal that increases in expectancy result in declining marginal increases in motivation.

In group 3, 26 working managers in a southeastern city completed a decision making exercise, with each subject's effort-level decision values providing operational measurements of the person's motivation to accomplish various projects within budgeted cost goals. Each subject provided eight decision values for each of three expectancy levels (0.10, 0.50, and 0.90). If, as proposed, increases in expectancy (expectation) result in declining marginal increases in motivational force, the differences between each subject's corresponding decision values when expectancy equaled 0.50 and when expectancy equaled 0.10 will be greater than the differences between each subject's corresponding decision values when expectancy equaled 0.90 and when it equaled 0.50. The results (df = 207, t = 4.1, $p < .001$) support the proposal that increases in expectancy result in declining marginal increases in motivation.

In group 4, 48 graduate students in engineering, management, or operations research at a midwestern university completed a decision making exercise, with each subject's effort-level decision values providing operational measurements of the person's motivation to earn an "A" in an academic course. Each subject provided eight decision values for each of three expectancy levels (0.00, 0.40, and 0.80). If, as proposed, increases in expectancy result in declining marginal increases in motivational force, the differences between each subject's corresponding decision values when expectancy equaled 0.40 and when it equaled 0.00 will be greater than the differences between each subject's corresponding decision values when expectancy equaled 0.80 and when expectancy equaled 0.40. The results ($df = 383$, $t = 5.2$, $p < .0001$) support the proposal that increases in expectancy result in declining marginal increases in motivation.

---

[2] An alternative approach for analyzing these data is to compute the average difference between the eight decisions each subject reached when expectancy equaled 0.30 and when expectancy equaled 0.00 and the average difference between the eight decisions each subject reached when expectancy equaled 0.30 and when expectancy equaled 0.60. The paired-samples t-test may then be used to compare these average differences. The study results presented in this paper are statistically significant if this alternative analysis approach is used.

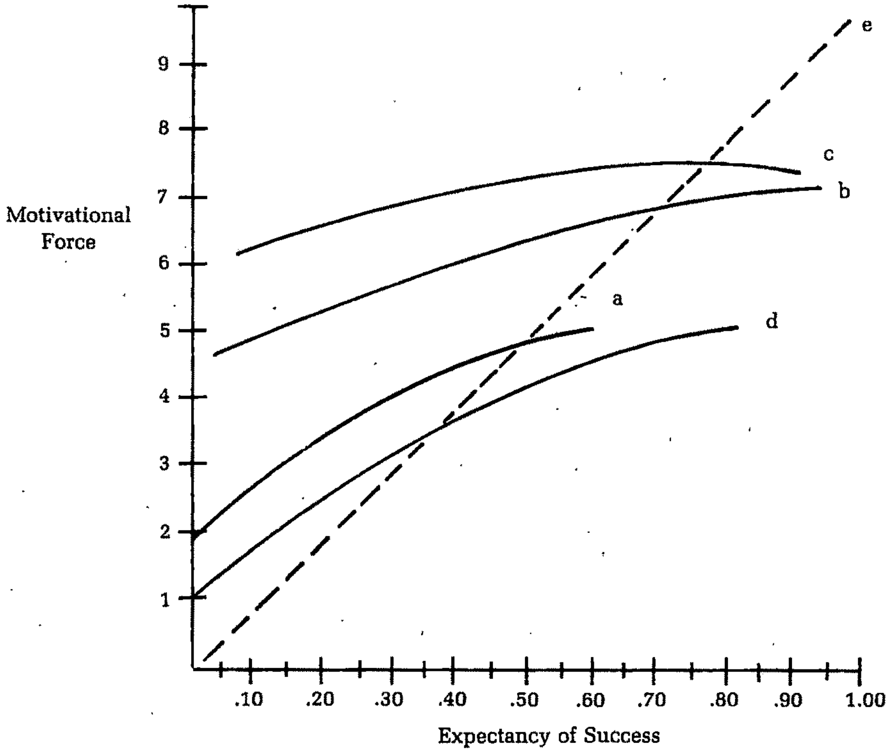
In group 5, 20 graduate students in management from the same midwestern university as group 4 came from completed the same decision making exercise, with each subject's effort-level decision values providing operational measurements of that person's motivation to earn an "A" in an academic course. Each person provided eight decision values for each of three expectancy levels (0.00, 0.20, and 0.60), with two of these expectancy levels differing from those of group 4. If, as proposed, increases in expectancy result in declining marginal increases in motivation, the differences between each subject's corresponding decision values when expectancy equaled 0.20 and when expectancy equaled 0.00 will be greater than one-half the differences between each subject's corresponding decision values when expectancy equaled 0.60 and when expectancy equaled 0.20. The results ($df = 159$, $t = 2.1$, $p < .02$) support the proposal that increases in expectancy result in declining marginal increases in motivation. Figure 1 summarizes the results for all five groups.

DISCUSSION

In a number of instances, the subjects' decisions differed from those implied by the expectancy force model, the expected utility model, or the prospect theory model. Since all three of these models contain the premise that individuals employ multiplicative information processing to arrive at their decisions, all three models imply that only a very small level of motivation should exist when the expectancy of success is nearly zero. They further imply that no motivation should exist when the expectancy of success is zero. However, as Figure 1 shows, many of the subjects' decisions indicated that a significant level of motivation existed when the expectancy of success was quite small (0.05 and 0.10) and that some positive level of motivation existed even when the expectancy of success was zero.

Previous research provides a possible explanation of these findings (Harrell & Stahl, 1984; Stahl & Harrell, 1981). Results of that research indicate that most of the subjects who provided our data employed additive information processing to arrive at their decisions, rather than the multiplicative information processing implied by the expectancy force model, the expected utility model, and the prospect theory model. If individuals employ

FIGURE 1
Expectancy and Motivational Force Relationships



[a] Group 1
[b] Group 2
[c] Group 3
[d] Groups 4 and 5
[e] Relationships implied by the expectancy force model

additive information processing to arrive at decisions concerning motivation, they should be expected to experience some level of motivation even if expectancy is very small or even zero when the valence of an outcome is large (Stahl & Harrell, 1981: 321). Since additive linear models usually explain nearly all of the variance in the decisions reached by most individuals even when their actual information processing is multiplicative (Harrell & Stahl, 1984; Stahl & Harrell, 1981; Slovic, Fishhoff, & Lichtenstein, 1977; Slovic & Lichtenstein, 1971), it seems possible the findings reported in this paper are typical of most people. Using this rationale, we might explain the empirical results provided by Kahneman and Tversky (1979: 265–273), who apparently assumed that their subjects employed the multiplicative information processing implied by the prospect theory model. If most of their subjects actually employed additive information processing and associated a large valence or utility with certain outcomes, we would expect them to have reached choices that neither the multiplicative expected utility model nor the multiplicative expectancy force model predict. The research of Matsui, Okada, and Mizuguchi (1981) appears to support this potential explanation for such behavior. They concluded that levels of motivation are, in some instances, more strongly influenced by the valence of an outcome than by the outcome's associated expectancy.

The data provided by all five groups of participants consistently support the proposal that increases in the expectancy of success result in declining marginal increases in individuals' levels of motivation. The study's results therefore provide straightforward support for Rynes and Lawler's (1983) suggestion about these relationships. Since the cases in the exercises used to gather the data involved positive nonzero outcomes, the results also provide support for the general relationships implied by the concave portion of Kahneman and Tversky's (1979) prospect theory value function.

CONCLUSIONS

To summarize, recent research on expectancy theory (Butler & Womer, 1985; Harrell & Stahl, 1984; Rynes & Lawler, 1983; Stahl & Harrell, 1981) has indicated that many, perhaps most, people employ additive information processing to arrive at decisions concerning motivation rather than the multiplicative information processing implied by the expectancy force model. This finding may account for the results presented in this study indicating that (1) many individuals exhibit a significant level of motivation even when their expectancy of success is very small or zero; and (2) increases in the expectancy of success result in declining marginal increases in motivation rather than in the proportionate increases implied by the expectancy force model.

Given these findings, a question might arise as to whether the expectancy force model, as it is commonly formulated and applied in motivation theory, is still useful. We believe that it is still useful as a conceptual model for explaining the factors that motivate individual behavior. For example, we believe it is correct to state that an individual's level of motivation is a function of the force model's valence and expectancy variables (Vroom, 1964). However, we believe the explanatory and predictive ability of the expectancy theory models would be substantially improved by integrating the empirical findings presented in the literature on behavioral decision theory into expectancy theory. Accordingly, we propose that some assumptions of prospect theory (Kahneman & Tversky, 1979; Tversky & Kahneman, 1981) be integrated into expectancy theory. These assumptions are that individuals (1) initially perform an editing activity during which they eliminate some alternatives and frame acts, outcomes, and contingencies, and (2) subsequently perform an evaluation activity during which they evaluate the remaining alternatives. We further propose that the expectancy valence model can effectively explain or predict the valence individuals associate with first-level outcomes after they have performed the initial editing activity. Moreover, we propose that the expectancy force model, modified to acknowledge that many individuals employ additive information processing, can be effectively employed to explain or predict the level of motivational force acting on an individual. As noted earlier, a large body of prior research indicates that additive linear models usually explain nearly all of the variance in the choices reached by most individuals even when their actual information processing is multiplicative (Harrell & Stahl, 1984; Stahl & Harrell, 1981; Slovic, Fishhoff, & Lichtenstein, 1977; Slovic & Lichtenstein, 1971). Accordingly, if an individual has performed the initial editing activity described by prospect theory (Kahneman & Tversky, 1979; Tversky & Kahneman, 1981), we would expect the additive form of the expectancy force model proposed in our earlier research (Stahl & Harrell, 1981: 321) will nearly always explain adequately the individual's choice behavior. We hope these proposals will stimulate other researchers to further examine these issues.

REFERENCES

Allais, P. 1953. The behavior of rational man in risk situations—A critique of the axioms and postulates of the American School. ***Econometrica***, 21: 503–546.

Arnold, H., & Evans, M. 1979. Testing multiplicative models does not require ratio scales. ***Organizational Behavior and Human Performance***, 24: 41–59.

Butler, J., & Womer, K. 1985. Hierarchical vs. non-nested tests for contrasting expectancy-valence models: Some effects of cognitive characteristics. ***Multivariate Behavioral Research***, 20: 335–352.

Cohen, J. 1978. Partialed products are interactions; partialed powers are curve components. ***Psychological Bulletin***, 85: 858–866.

Cohen, J., & Cohen, P. 1975. ***Applied multiple regression/correlation analysis for the behavioral sciences***. New York: John Wiley & Sons.

Connolly, T. 1976. Some conceptual and methodological issues in expectancy models of work performance motivation. ***Academy of Management Review***, 1(4): 37–47.

Coombs, C. 1975. Portfolio theory and the measurement of risk. In M. Kaplan & S. Schwartz (Eds.), ***Human judgment and decision processes***: 64–83. New York: Academic Press.

Ellsberg, D. 1961. Risk, ambiguity, and the Savage axioms. ***Quarterly Journal of Economics***, 75: 643–649.

Harrell, A., & Stahl, M. 1984. Modeling managers' effort-level decisions for a within-persons examination of expectancy theory in a budget setting. ***Decision Sciences***, 15: 52–73.

Heneman, H., & Schwab, D. 1972. Evaluation of research on expectancy theory predictions of employee performance. ***Psychological Bulletin***, 78: 1–9.

House, R., Shapiro, H., & Wahba, M. 1974. Expectancy theory as a predictor of work behavior and attitude: A reevaluation of empirical evidence. ***Decision Sciences***, 5: 481–506.

Kahneman, D., & Tversky, A. 1979. Prospect theory: An analysis of decision under risk. ***Econometrica***, 47: 263–291.

Keeney, R., & Raiffa, H. 1976. ***Decisions with multiple objectives: Preferences and tradeoffs***. New York: John Wiley & Sons.

Kopelman, R. 1977. Across-individual, within-individual and return on effort versions of expectancy theory. ***Decision Sciences***, 8: 651–662.

Locke, E. 1975. Personal attitudes and motivation. ***Annual Review of Psychology***, 25: 457–480.

Matsui, T., Okada, A., & Mizuguchi, R. 1981. Expectancy theory predictions of the goal theory postulate, "The harder the goals, the higher the performance." ***Journal of Applied Psychology***, 66: 54–58.

Mitchell, T. 1974. Expectancy models of job satisfaction, occupational preference, and effort: A theoretical, methodological, and empirical appraisal. ***Psychological Bulletin***, 81: 1053–1077.

Mitchell, T. 1979. Organizational behavior. ***Annual Review of Psychology***, 30: 243–281.

Mitchell, T., & Beach, L. 1977. Expectancy theory, decision theory, and occupational preference and choice. In M. F. Kaplan & S. Schwartz (Eds.), ***Human judgment and decision processes in applied settings***: 203–226. New York: Academic Press.

Rynes, S., & Lawler, J. 1983. A policy-capturing investigation of the role of expectancies in decisions to pursue job alternatives. ***Journal of Applied Psychology***, 68: 620–631.

Samuelson, P. 1967. ***Economics***. New York: McGraw-Hill Book Co.

Schmidt, F. 1973. Implications of a measurement problem for expectancy theory research. ***Organizational Behavior and Human Performance***, 10: 243–251.

Schwab, D., Olian-Gottlieb, J., & Heneman, H., III. 1979. Between-subjects expectancy theory research: A statistical review of studies predicting effort and performance. ***Psychological Bulletin***, 86: 139–147.

Slovic, P., Fishhoff, B., & Lichtenstein, S. 1977. Behavioral decision theory. ***Annual Review of Psychology***, 28: 1–39.

;Slovic, P., & Lichtenstein, S. 1971. Comparison of Bayesian and regression approaches to the study of information processing in judgment. ***Organizational Behavior and Human Performance***, 6: 649–744.

Stahl, M., & Harrell, A. 1981. Modeling effort decisions with behavioral decision theory: Toward an individual differences model of expectancy theory. ***Organizational Behavior and Human Performance***, 27: 303–325.

Stahl, M., & Harrell, A. 1983. Using decision modeling to measure second level valences in expectancy theory. ***Organizational Behavior and Human Performance***, 32: 23–34.

Tversky, A., & Kahneman, D. 1981. The framing of decisions and the psychology of choice. ***Science***, 211: 453–458.

Vroom, V. 1964. ***Work and motivation***. New York: John Wiley & Sons.

Wahba, M., & House, R. 1974. Expectancy theory in work and motivation: Some logical and methodological issues. ***Human Relations***, 27(2): 121–147.

Wolf, G., & Connolly, T. 1981. Between-subject designs in testing expectancy models: A methodological note. ***Decision Sciences***, 12: 39–45.

Zedeck, S. 1977. An information processing model and approach to the study of motivation. ***Organizational Behavior and Human Performance***, 18: 47–77.

Adrian Harrell earned his Ph. D. degree at the University of Texas; he is a professor of accounting at the University of South Carolina. His current research interests include using McClelland's trichotomy of needs theory to predict the future job satisfaction and voluntary turnover of public accounting professionals.

Michael Stahl earned his Ph. D. degree at the Rensselaer Polytechnic Institute; he is a professor of management and Head of the Department of Management at Clemson University. His current research interests include using judgment modeling to examine the descriptive validity of normative financial decision models.

CALL FOR PAPERS
FOR THE 1987 MEETINGS OF THE SOUTHWEST ACADEMY OF MANAGEMENT

The 1987 meetings will be held March 10-14, at the Hyatt Regency, Houston, Texas. Submissions of competitive papers and symposium proposals must be postmarked by September 5, 1986.

For more information or submissions contact:

C. Aaron Kelley—(502) 588-7830
Department of Management
School of Business
University of Louisville
Louisville, Kentucky 40292

Please send all address corrections to:
WALTER B. NEWSOM
Secretary - Treasurer
ACADEMY OF MANAGEMENT
P.O. DRAWER KZ
Mississippi State, MS 39762

New From Praeger . . .

GUIDES TO SUCCESS IN INTERNATIONAL BUSINESS

For Policymakers and Moneymakers

SERVICE-LED GROWTH
The Role of the Service Sector in World Development
Dorothy I. Riddle, American Graduate School of International Management

Are services a post-industrial phenomenon? Does growth in the service sector result in lower productivity? These and other questions are addressed in this groundbreaking study of over 80 countries at four levels of economic development that concludes that developing domestic services and services exports may be more important to developing nations than even in industrialized countries.

1986 289 pp. LC 85-16743
Cloth ISBN 0-275-92041-0
$27.95

THE JAPANESE SCHOOL
Lessons for Industrial America
Benjamin Duke, International Christian University, Japan
Forewords by Clark Kerr and James M. Hestor
With an Introduction by Edwin Reischauer

"one of the most important books on Japanese society."
—Peter Drucker

Does Japan's competitive industrial edge begin with its schools? This book explores the impact of Japan's school system on its economic development since WW II, and demonstrates how the U.S. can increase its competitiveness by strengthening American schools.

1986 208 pp.
Cloth ISBN 0-275-92109-3
$32.95
Paper ISBN 0-275-92003-8
$12.95

THE INTERNATIONAL BUSINESSWOMAN
A Guide to Success in the Global Marketplace
Marlene L. Rossman, Rossman Graham Associates

A nuts-and-bolts guide for women doing business abroad. Time-tested tips on careers, negotiating strategies, cultural idiosyncrasies, etc.

1986 approx. 160 pp.
Cloth ISBN 0-275-92000-3
$32.95
Paper ISBN 0-275-92001-1
$10.95

For your copy of these or other Praeger business titles, check your local bookstore, or write:

Praeger PUBLISHERS
A Division of Greenwood Press, Inc.
88 Post Road West — Box 5007
Westport, CT 06881

State University of New York Press
announces a new series in
LEADERSHIP
edited by
Barbara Kellerman
Institute for Leadership Studies
Fairleigh Dickenson University

The Press is inviting manuscripts on a wide range of leadership topics. These topics may pertain to: the public or private sector, domestic, cross-cultural, or international perspectives, and individual leaders or leadership in small groups, organizations and large collectives. To attract a wide general readership and facilitate course adoptions, books in the series will appear simultaneously in cloth and paper editions.

All inquiries, manuscripts, and book proposals submit to:

Michele Martin, Editor
SUNY Press
State University Plaza
Albany, New York 12246-0001

Introducing

Salenger films/videos in cooperation with
The Academy of Management
presents a new film

A Question of Mangement:
a historical perspective

Celebrating the Centennial of Management
and the 50th anniversary of The Academy of Management

How can productivity be increased? What can a manager do to get the most work in the least time? With that question, the science of management was born. This epoch making film reviews the history of management theory and practice—from its earliest beginnings, through Frederick Taylor and scientific management, to todays' Systems Theory and Contingency Approach. Here are the answers to many of the problems that confront todays' managers and supervisors.

—35 MINUTES

LEASE / PURCHASE $595 **RENTAL $125** **PREVIEW $35**

Name / Title ______________________

Institution ______________________

Address ______________________

City, State, Zip ______________________

Phone ______________________

A814

SALENGER FILMS/VIDEOS 1635—12th St., SANTA MONICA, CA 90404 • (213) 450-1300

Academy of Management Placement Roster 1986-87 Advertising Rates and Data

THE PLACEMENT ROSTER AND YOU

Advertising in the Academy of Management Placement Roster gives you an unequaled opportunity to reach a unique market. Published each fall and spring, the Placement Roster provides services for members and nonmembers, for those applying for positions as well as for those advertising available positions. Placement listings include faculty teaching positions, departmental chairs, administrative positions, and public and private sector positions. A display ad along with an institutional listing greatly enhances the visibility of any open position.

DISPLAY ADVERTISING RATES

The following dimensions and rates are available for display advertising in the Placement Roster. Listed rates are for camera ready copy only. We do offer you production services but at additional cost. Such services include typesetting, art, borders, photos, etc. Color also is available at $175 for each additional color.

Black and White	Size (Inches)	Rate
Full Page	4½ × 7¼	$350
Two Page Spread		650
Half Page	4½ × 3⅝	200
Inside Back Cover		450
Each Additional Color		175

RESTRICTIONS

The following restrictions apply to all display advertising:

1. No bleeds
2. Inside Front Cover, Front, Back and First Page Not Available.
3. Screens are 100 line dot prints.

IMPORTANT DATE

Issue Date:	Copy/Insertion Order Deadline:
November 11, 1986	October 10, 1986
April 13, 1987	March 13, 1987

FOR MORE INFORMATION

Please send all correspondence to:

Jim Townsend
D/MGT, Calvin Hall, Kansas State University
Manhattan, Kan. 66506
Phone: (913) 532-6050

Announcing the Celebration
of the

CENTENNIAL OF MANAGEMENT
and
THE 50th ANNIVERSARY OF THE ACADEMY OF MANAGEMENT

1986 National Academy of Management Meeting
August 13-16, 1986
Pre-Convention Activities/Meetings
August 11 (12 noon) to August 13 (12 noon)

Hyatt Regency-Downtown
Chicago, Illinois

For further information contact:

Don Hellriegel
Vice President-Program Chair
Department of Management
Texas A & M University
College Station, Texas 77843
(409) 845-1724

Applications for membership to the Academy may be obtained from:

Joseph P. Yaney
Director of Membership
College of Business Administration
Texas Tech University
Lubbock, Texas 79409-4320